

ESCUELA TECNICA SUPERIOR DE INGENIEROS DE TELECOMUNICACION

APUNTES DE
ESTEREOTRONICA.
QUINTO CURSO

JOSE A. MARTIN PEREDA
Profesor de la Asignatura

CURSO 1.967-68

ESTEREOEOTRONICA

I. INTRODUCCION

I. 1. - Naturaleza de la Radiación

Uno de los hechos más claros acerca de la naturaleza de la luz, es que la sombra producida por los cuerpos opacos, se propaga en línea recta, es decir la luz viaja según esta trayectoria. Este hecho condujo a Newton a formular su "Teoría Corpuscular de la luz", según la cual la energía fluye de la fuente luminosa en forma de energía cinética de una corriente de partículas. El fenómeno de interferencias no pudo ser explicado por esta teoría, por lo que fué reemplazada por la teoría ondulatoria.

El fenómeno de interferencias se representa en forma esquemática en la Figura 1. S es una fuente luminosa que está separada de la pantalla CD por otra opaca EF en la que se han practicado dos pequeños agujeros A y B, de tal manera que $\overline{SA} = \overline{SB}$. Estos agujeros actuarán pues como nuevas fuentes de luz. Si la luz fuera corpúscular la iluminación en cualquier punto de la pantalla CD sería siempre mayor cuando los dos agujeros estuvieran abiertos que cuando se cierra uno de ellos. Por otra parte, si la luz se propaga de manera ondulatoria no ocurrirá lo mismo; para un punto P en el que se cumpliera que $\overline{PA} - \overline{PB} = 0$ ó múltiplo de una λ sumaran los efectos de ambas fuentes secundarias, pero si $\overline{PA} - \overline{PB} = \frac{n}{2} \lambda$ los efectos de estas fuentes se restarán (interferencia destructiva).

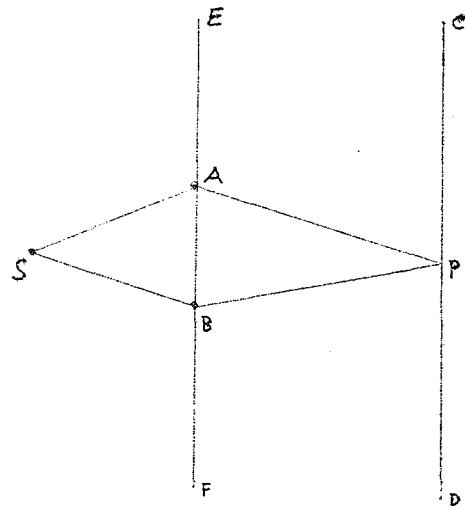


Fig. 1: Experimento de interferencia.

Por otra parte, ciertos fenómenos tales como el aporte de energía a la materia por la luz y el estudio de las radiaciones de muy corta longitud de onda, luz y rayos X, no pueden ser explicadas mediante la teoría ondulatoria, y sin embargo lo son fácilmente la teoría corpuscular. Esta dualidad es esencial para la comprensión de la mecánica cuántica y vamos a considerarla en algún detalle.

Aparte del ojo humano, las formas por las cuales la luz y las radiaciones de muy corta λ pueden ser observadas y medidas son:

- 1.- Fotografía
- 2.- Foto conductividad
- 3.- Efecto fotográfico. ?

Veamos algunos casos. Einstein demostró que cuando incide una radiación de frecuencia ν_0 sobre un campo fotoemisor la energía suministrada es un múltiplo de $h \nu_0$. En la fotoconductividad se comprueba que hay una frecuencia límite por debajo de la cual no se produce el fenómeno. En la Fotografía se comprueban efectos idénticos. La existencia de un límite inferior es común a todos los efectos fotoquímicos y fotoeléctricos y en todos los que hay un proceso de comunicación de energía entre luz y materia, a excepción de la inducción de corriente en conductores. En este último caso no hay frecuencia límite inferior, como es bien conocido en el estudio de antenas y propagación de ondas, gracias al Sr. Huertas. Esta aparente separación de la teoría cuántica está estrechamente relacionada con la existencia de electrones libres en los metales.

Los efectos estudiados demuestran que cuando la luz cede energía a la materia se comporta como si fuera un haz de partículas, cada una de ellas con una energía $h \nu$.

Esta dualidad en la naturaleza de la luz, por otra parte totalmente inesperada, ha penetrado en el campo de la Física Teórica por el camino experimental, y no ha sido nunca "explicada". Todavía más, esta naturaleza dual no está restringida unicamente a la luz, sino que es común a la mecánica de las pequeñas partículas, átomos electrones etc. Debe ser aceptada, entonces como una ley fundamental de la naturaleza en la cual debe basarse cualquier esquema formalmente válido que pretenda explicar la mecánica

I. 2.- Integral de fase. Aplicacion

En un intento de explicación de ciertos rayos espectrales, Sormerfeld, haciendo uso de la teoría de Bohr y de Plank, postuló que:

$$\oint \vec{P} \cdot d\vec{s} = nh \quad (1)$$

donde $\vec{P} = m\vec{v}$ es la cantidad de movimiento o momentum y la integral curvilinea está extendida a todo el recorrido de la partícula.

Veamos con un ejemplo a donde conduce este postulado.

Consideremos un electrón de masa m y carga e dando vueltas alrededor de un núcleo de hidrógeno. La fuerza culombiana de atracción será

$$F = \frac{e^2}{4 \pi \epsilon_0 r^2} \quad (2)$$

En equilibrio dinámico dicha fuerza debe ser compensada por la fuerza centrífuga del electrón

$$\frac{e^2}{4 \pi \epsilon_0 r} = \frac{m v^2}{r} \quad (3)$$

de esta igualdad podrán deducirse las expresiones de:

Energía cinética

$$T = \frac{1}{2} m v^2 \quad (4)$$

$$T = \frac{1}{2} \frac{e^2}{4 \pi \epsilon_0 r}$$

Energía potencial

$$V = \int_{\infty}^r F dr$$

$$V = \int_{\infty}^r \frac{e^2}{4 \pi \epsilon_0 r^2} dr \quad (5)$$

$$V = - \frac{e^2}{4 \pi \epsilon_0 r}$$

luego

$$V = -2T \quad (6)$$

Momento angular.

$$C = m v r \quad (7)$$

$$C = \sqrt{\frac{e^2 m r}{4 \pi \epsilon_0}}$$

Aplicando ahora el postulado de Sommerfeld:

$$\oint m v ds = 2 \pi m v r = 2 \pi \sqrt{\frac{e^2 m r}{4 \pi \epsilon_0}} = 2 \pi C = n h \quad (8)$$

Calculemos ahora la Energía Total E

$$E = T + V = - T = - \frac{2 \pi^2 m e^4}{(4 \pi \epsilon_0)^2 h^2} \cdot \frac{1}{n^2} \quad (9)$$

el radio

$$r = \frac{n^2 h^2 4 \pi \epsilon_0}{4 \pi^2 m e^2} \quad (10)$$

y el momento angular

$$C = \frac{n h}{2 \pi} \quad (11)$$

Es muy frecuente en mecánica cuántica el uso de

$$\hbar = h/2 \pi \quad (12)$$

M E C A N I C A O N D U L A T O R I A

1.- COMPORTAMIENTO CUANTICO

1-1.- Mecánica Atómica

9 experimentos

La mecánica cuántica es la descripción del comportamiento de la materia y la luz en todos sus aspectos, y en particular en la escala atómica. Los objetos en esta escala tan pequeña no se conducen como ninguna cosa de la que podemos tener experiencia directa.

Newton supuso que la luz estaba formada por partículas, pero más tarde se descubrió que se comportaba como una onda. Sin embargo, al comienzo del siglo XX, se redescubrió que algunas veces se comportaba como una partícula. Lo mismo sucede con el electrón. En realidad no se comportan ni como partículas ni como ondas.

Pero hay una semejanza entre todos los objetos de tamaño atómico. Los electrones, protonos, neutrones, fotones etc. se comportan como "ondas-partículas" o como se les llame. La única forma de intentar aprender algo acerca de estas partículas es realizar una serie de suposiciones abstractas e imaginarias, completamente desligadas de nuestra experiencia directa.

Vamos a intentar llegar inmediatamente al corazón de esta conducta misteriosa, presentando un fenómeno que nos resulta completamente imposible explicar por cualquier método clásico, y que contiene en realidad el único misterio de la mecánica cuántica. No podemos hacer que este misterio desaparezca explicando los hechos. Pero haciendolo así, habremos mostrado las peculiaridades básicas de la mecánica cuántica.

1-2.- Un experimento con balas.

Vamos a intentar comprender el comportamiento de los electrones, comparándolo y contrastándolo, en un conjunto de experimentos, con el comportamiento más familiar de partículas como balas, y ondas, como ondas de agua. Realizemos primero el siguiente experimento (Fig. 1.1)

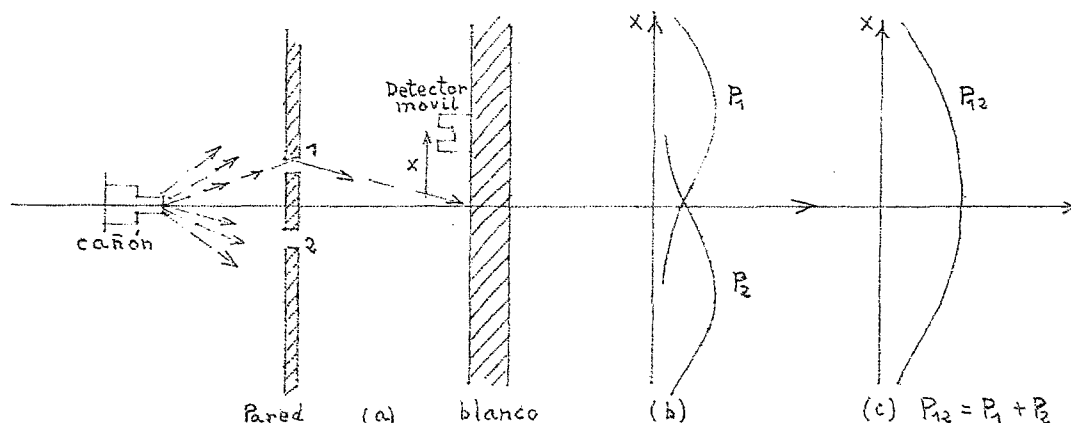


Fig. 1.1.- Experimento de interferencia con balas

Un cañón dispara un haz de balas. Si el cañón no es muy bueno - este haz será suficientemente ancho. Delante del cañón hay una pared con dos orificios lo suficientemente grandes para que pase una sola bala. Más allá de la pared está un muro que sirve de blanco, que "absorberá" las balas que choquen contra él. En la cara anterior de este muro colocamos una caja conteniendo arena, que pasará y almacenará cualquier bala que penetre en él, por lo tanto actuará como un "detector" de balas. Este detector puede moverse hacia arriba y hacia abajo, o sea, en la dirección del eje x .

Cuál será la probabilidad de que una bala que atraviesa los orificios de la pared, llegue al blanco a una distancia x del centro?

Si suponemos que el cañón dispara al mismo ritmo durante todo el experimento manteniendo el detector durante un intervalo de tiempo estándar a una distancia x del centro, el número de balas recogidas será proporcional a la probabilidad pedida. Realizando medidas en distintas posiciones podremos ir contestando a la pregunta.

Para lo que nos proponemos deberemos imaginar un experimento idealizado, en el que las balas son indestructibles. Así encontraremos que las balas llegan siempre al blanco en partes enteras que llamaremos "corpúsculos" y en el detector siempre contaremos un número entero de balas. Si el cañón dispara a un ritmo muy lento, en un momento dado llegará al blanco una sola bala o ninguna. También, el tamaño del "corpúsculo" no depende del ritmo de fuego: "Las balas llegan siempre en corpúsculos idénticos". Con nuestro detector mediremos la probabilidad de llegada de un corpúsculo, como función de x . El resultado de la medida lo dibujamos en la parte (c) de la figura 1.1. La llamamos P_{12} por que las balas pueden llegar atravesando el orificio 1, o bien el 2. No nos sorprende la forma de la figura, pero podríamos preguntarnos por que P_{12} tiene un máximo para $x = 0$. Podríamos entender este hecho si repetimos el experimento, primero tapando el orificio 2, y luego tapando el 1. Cuando tapamos 2, las balas pasan a través de 1 y obtenemos la curva P_1 de (b). El máximo de P_1 estará en el valor de x alineado con el cañón y el orificio 1. Cuando cerramos 1, encontraremos la curva simétrica P_2 dibujada en (b). Comparando (b) y (c) encontramos que

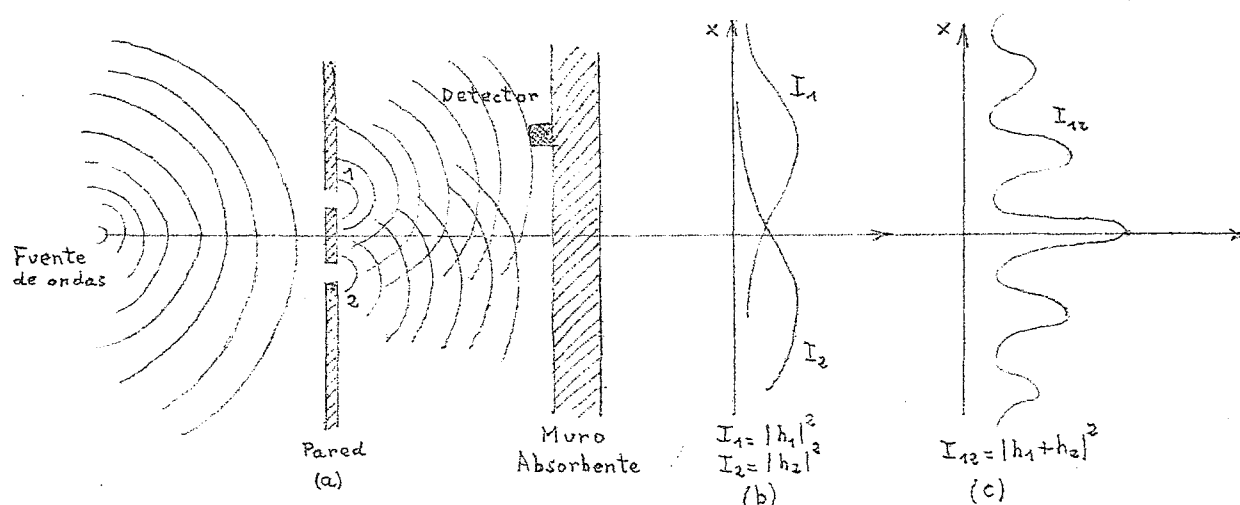
$$P_{12} = P_1 + P_2 \quad (1.1)$$

Las probabilidades se suman. El efecto con ambos orificios abiertos es la suma de los efectos con un único orificio abierto. Esta observación la llamaremos de "no interferencia" por una razón que se verá más tarde. Resumiendo: Las balas llegan en corpúsculos y su probabilidad de llegada no presenta inter

ferencia.

1-3.- Un experimento con ondas.

Realizaremos ahora un experimento con ondas de agua, con el dispositivo de la figura 1.2.



La "fuente de ondas" es un pequeño objeto que vibra en la superficie del agua, arriba y abajo movido por un motor, produciendo ondas circulares. Tenemos también la pared del experimento anterior y luego un muro absorbente para que no produzca ondas reflejadas (P. ej. una playa de arena de pequeña inclinación). El detector es un aparato que mide la "intensidad" de la onda. Puede ser un aparato medidor de la altura del movimiento ondulatorio, pero en una escala graduada proporcionalmente al cuadrado de dicha altura.

Lo primero que descubrimos con este aparato, es que la intensidad puede tener cualquier valor. Si la fuente se mueve muy poco, en el detector se registrará un movimiento ondulatorio muy pequeño, pero cuanto mayor sea el movimiento de la fuente, mayor será la intensidad en el detector. No podremos decir que hay "corpuscularización" en la intensidad de onda.

Realizando las medidas para diversos valores de x , obtendremos la interesante curva I_{12} en la parte (c) de la figura 1.2.

Este es el típico modelo de interferencia con ondas que ya se conoce. La onda original se difracta en los orificios y nuevas ondas circulares parten de cada uno de ellos. Si cubrimos un hueco cada vez y medimos la distribución de intensidades, encontramos las curvas dibujadas en la parte (b) de la figura. I_1 es la intensidad de la onda que parte de 1 (con 2 cerrado) e I_2 la de 2 (con 1 cerrado).

La intensidad I_{12} no es la suma de I_1 e I_2 . Diremos que hay in-

terferencia de las dos ondas. En los lugares en que la curva I_{12} es máxima, las dos ondas están en fase, se suman los picos y se produce una gran intensidad. Diremos que se "interfieren constructivamente" en tales sitios. Habrá interferencia constructiva cuando la distancia del detector a un orificio es un número entero de longitudes de onda más larga (ó más corta) que la distancia de otro orificio.

Cuando las dos ondas llegan al detector con una diferencia de fase de π , la amplitud resultante es la diferencia de las dos amplitudes. Las ondas se "interfieren destructivamente" y hay un bajo valor de la intensidad (diferencia de distancias desde el detector a los orificios = número impar de medias longitudes de onda).

Podemos expresar la relación cuantitativa de este modo. Para el orificio 1, la altura de la onda puede ser descrita como la parte real de $h_1 e^{j\omega t}$, donde la amplitud h , es en general un número complejo. La intensidad es proporcional al cuadrado del valor absoluto de la amplitud $|h_1|^2$. Análogamente para 2 será $h_2 e^{j\omega t}$ y $|h_2|^2$. Cuando los dos huecos están abiertos, las ondas se sumarán y el valor será $(h_1 + h_2)e^{j\omega t}$ y la intensidad $|h_1 + h_2|^2$. Omitiendo las constante de proporcionalidad tendremos.

$$I_1 = |h_1|^2 \quad I_2 = |h_2|^2 \quad I_{12} = |h_1 + h_2|^2 \quad (1.2.)$$

$$|h_1 + h_2|^2 = |h_1|^2 + |h_2|^2 + 2|h_1||h_2| \cos \delta \quad (1.3.)$$

donde δ es la diferencia de fase entre h_1 y h_2

$$I_{12} = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos \delta \quad (1.4.)$$

y este es resultado observado en la parte (c) de la figura.

1-4.- Un experimento con electrones.

Imaginemos ahora un experimento similar con electrones (fig.1.3)

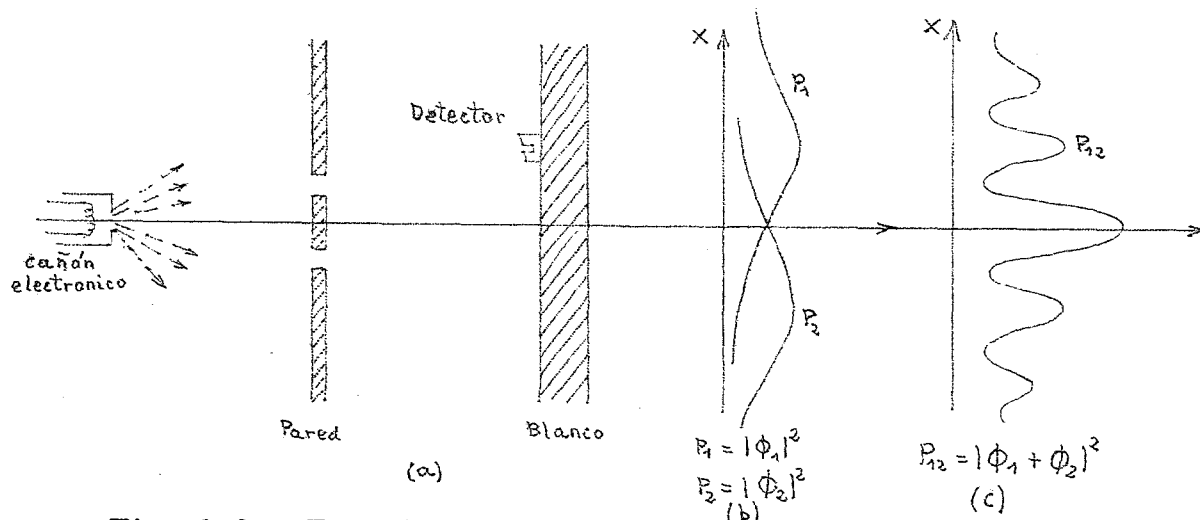


Fig. 1.3.- Experimento de interferencia con electrones.

Todos los electrones que salen del cañón tienen aproximadamente la misma energía. El detector puede ser un contador geiger o, quizás mejor, un multiplicador electrónico conectado a un altavoz.

Señalemos que tal experimento no se puede realizar de esta forma, porque la escala del aparato debería ser increíblemente pequeña. Es un experimento "pensado". Conocemos de otros experimentos los resultados que se de rían obtener y hemos escogido estas escalas y proporciones para mayor faci lidad.

Lo primero que notamos, es que oímos fuerte "clicks" procedentes del detector. Y todos los "clicks" son el mismo. No hay "medios clicks".

También notaremos que los clicks se oyen muy erráticamente. Si contamos los clicks del llegada durante un tiempo suficientemente largo -diga-- mos durante varios minutos, y volvemos luego a contarlos para otro periodo - igual, encontraremos que los dos números son aproximadamente el mismo. - Podremos hablar de "razón media" a la cual se oyen los clicks.

Si movemos el detector arriba y abajo, la velocidad de oída de los clicks será más rápida o más lenta, pero la intensidad o tamaño de cada click es siempre la misma. Si bajamos la temperatura del filamento, la velocidad de los clicks bajará pero su intensidad seguirá siendo la misma. También notaríamos que si ponemos dos detectores en el blanco, uno u otro darán un click pero nunca ambos a la vez. De todo esto conduciremos que los electrones llegan al blanco en "corpúsculos", todos del mismo tamaño. "Los electrones llegan siempre en corpúsculos idénticos.

Cuál es la probabilidad de que un corpúsculo electrónico llegue - al blanco a varias distancias x del centro?" Procediendo como con las balas nos resulta la curva marcada con P_{12} en la parte (c) de la figura 1.3. Exac- tamente y aunque parezca mentira esto es lo que los electrones hacen.

1-5.- La interferencia de las ondas electrónicas.

Vamos ahora a intentar entender el comportamiento de los elec- trones analizando la curva de la figura 1.3. Formulemos primero la siguiente

~~"Proporción"~~ :

Proporción

Proporción A: Cada electrón pasa, o bien a través del hueco 2, o bien a través del hueco 1.

Suponiendo que esta proporción es cierta, todos los electrones que llegan al blanco se pueden dividir en dos clases : (1) Los que pasan a través - del orificio 1, y (2) los que lo hacen por 2. De esta forma la curva observada debe ser la suma de los efectos de los electrones que atraviesan por 1 y de -

los que pasan por 2. Comprobemoslo experimentalmente. Igual que hicimos - en el experimento con balan, bloqueando 2 y contando los clicks del detector - obtenemos P_1 , de modo simétrico obtenemos P_2 (figura 1.3 (b)).

Se ve claramente que P_{12} no es la suma de P_1 y P_2 . Entonces - diremos "Hay interferencia".

Para electrones : $P_{12} \neq P_1 + P_2$

Cómo puede producirse tal interferencia? Quizás podríamos decir:

'Bien, esto significa, presumiblemente, que no es verdad que los electrones o corpúsculos vayan a través de 1 o de 2, ya que si así sucediera, las proba**bilidades** deberían sumarse. Quizás siguen un camino más complicado. Se par**te** por la mitad y" Pero no! Esto no puede ser porque llegan siempre - en corpúsculos....." Bien, quizás alguna pasa por 1, después regresa por 2, y después de unas cuantas vueltas más, o va por otro camino complicado.... entonces, al cerrar 2, hemos alterado la posibilidad de que un electrón que comenzó pasando por 1, llegase al blanco....." Pero cuidado! Hay algunos - puntos a los que llegan muy pocos electrones cuando ambos orificios están - abiertos, pero que reciben muchos si cerramos un hueco. Sin embargo, en el centro del gráfico, P_{12} es mayor que el doble de $P_1 + P_2$. Es como si pensa**ramos** que cerrando un hueco decrece el número de los electrones que atravie**san** al otro. Parece muy difícil explicar ambos efectos proponiendo que los - electrones viajen por caminos complicados.

Todo es muy misterioso. Y cuanto más se lo mire más misterio**so** parece. Se ha intentado por muchos medios explicar la curva P_{12} en termi**no** de los electrones individuales dando vueltas por caminos complicados. Pe**ro** nadie ha alcanzado el éxito.

Pero, aunque sea sorprendente, la relación matemática entre - P_{12} y P_1 y P_2 es extremadamente sencilla. Porque P_{12} es casi igual a la cur**va** I_{12} de la fig. 1.2. y aquella fué sencilla. Lo que está llegando al blanco - puede ser descrito por dos números complejos que podemos llamar Φ_1 y Φ_2 - (desde luego son funciones de x). El cuadrado del valor absoluto de Φ_1 - nos dá el efecto cuando solo está abierto 1. Esto es $P_1 = |\Phi_1|^2$. Analogamen**te** $P_2 = |\Phi_2|^2$. Y el efecto combinado de los dos huecos es igual a $P_{12} = |\Phi_1 + \Phi_2|^2$. Son las mismas matemáticas que en el caso de las ondas acuá**ti**cas.

Observemos la siguiente conclusión: Los electrones llegan en cor**púsculos** o paquetes, como las partículas, y la probabilidad de llega**da** de esos corpúsculos se distribuye como la distribución de intesidad de una onda. For

eso se dice que un electrón se "comporta" algunas veces como una partícula y algunas veces como una onda.

Debemos señalar que en mecánica cuántica estas amplitudes serán representadas por números complejos, no por sus partes reales solamente, como en las ondas de agua.

Hay muchas sutilizas encerradas en el hecho de que la naturaleza obra de este modo. Por ejemplo, ya que el número de las que llegan a un punto particular no es igual al número de los que llegan a través de 1 más el de los que llegan a través de 2, como se seguiría de la proporción, indudablemente concluiremos que la proporción A es falsa. No es verdad que cada electrón vaya o a través de 1 o a través de 2. Pero esta condición puede probarse por otro experimento.

1-6.- Vigilando los electrones.

Realizemos ahora el siguiente experimento. Añadimos a nuestro aparato electrónico una fuente de luz muy fuerte, situada detrás de la pared y entre los dos huecos, como muestra la fig. 1.4. Sabemos que las cargas eléctricas reflejan la luz. Cuando un electrón pasa, si es que pasa, en su ca-

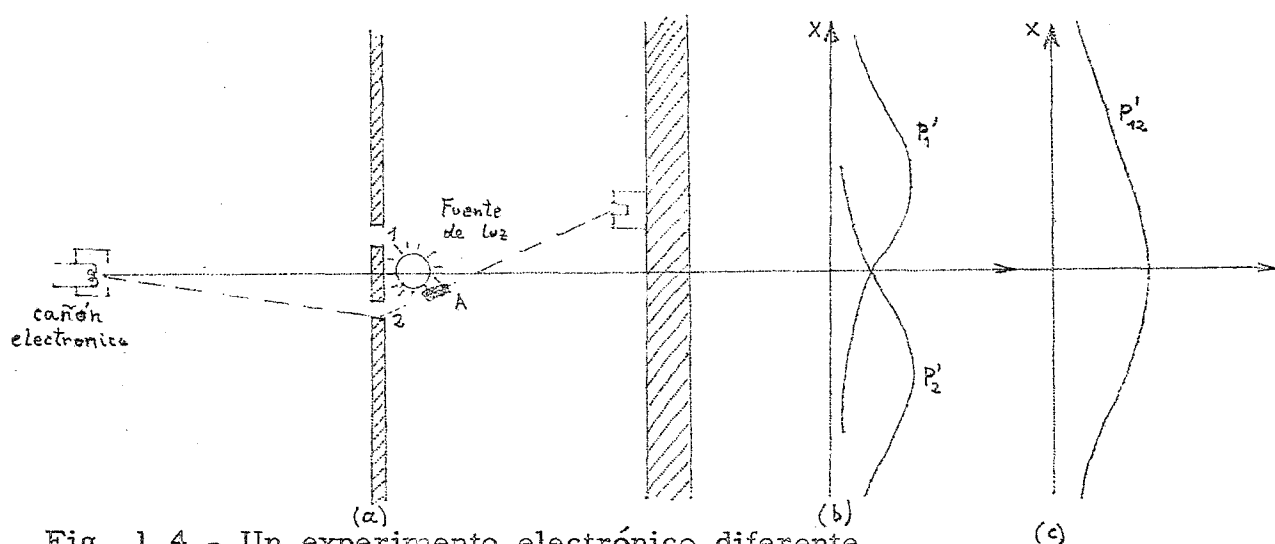


Fig. 1.4.- Un experimento electrónico diferente.

mino hacia el detector, reflejará algo de luz hacia nuestro ojo, y podremos ver por donde va el electrón. Si fuera por 2, como en la figura, veríamos un flash de luz proveniente del punto A. Si el electrón pasará por 1, el flash provendrá de la vecindad del hueco superior. Si vieramos luz al mismo tiempo por los dos huecos es que el electrón se divide por la mitad. Hagamos el experimento.

Lo que observamos es lo siguiente. Cada vez que oímos un click en el detector, veremos un flash luminoso cerca del hueco 1 o del 2, pero nunca en ambos a la vez, no importa donde pongamos el detector. O sea que

cuando vigilamos a los electrones encontramos que van a través de un hueco, o a través del otro. Experimentalmente, la proporción A es necesariamente cierta.

Esto se contradice con lo anterior. Por qué no es $P_{12} = P_1 + P_2$? Sigamos con el experimento para cada posición del detector, realizando previamente un cuadro con dos columnas, cuando oigamos un "click" apuntaremos un punto a la columna 1 si vemos el flash cerca del hueco 1, y análogamente en la columna 2, si es cerca de 2. Todos los electrones que llegan al detector se puede dividir en dos clases, los que pasan por 1, y los que pasan por 2. Del número de puntos en la columna 1 obtenemos la probabilidad P'_1 de que un electrón llegue al detector via hueco 1, y análogamente con la columna 2 la probabilidad de P'_2 de que un electrón llegue al detector via hueco 2. Si repetimos las medidas para muchos valores de x obtenemos las curvas para P'_1 y P'_2 de (b). Vemos también que son similares a las obtenidas en el experimento anterior, lo que no es sorprendente, los electrones se distribuyen por ahora lo mismo cuando los vigilamos que cuando no los vigilamos.

Pero esperemos! Qué nos ha pasado con la probabilidad total la probabilidad de que el electrón llegue al detector por cualquier camino? Tenemos ya la información. Supongamos que nunca hemos mirado la luz de los flashes y hemos contado los corpúsculos por los clicks del detector, que hemos separado en dos columnas. Luego debemos sumar estos números y encontraremos que $P'_{12} = P'_1 + P'_2$. Cuando vigilamos los electrones, desaparece la vieja curva de interferencia P_{12} y tenemos una nueva P'_{12} que no muestra interferencia. Si apagamos la luz volvemos a obtener P_{12} .

Debemos concluir que cuando observamos los electrones su distribución sobre la pantalla es distinta de cuando no los miramos. Quizás es nuestra fuente de luz la que les perturba?. Debe ser que la luz que se refleja en ellos debe empujarlos bruscamente y perturba su movimiento. Efectivamente, el campo eléctrico de la luz actúa sobre la carga ejerciendo una fuerza sobre ella. Al tratar de vigilar los electrones, hemos cambiado su movimiento. Estos efectos de los fotones al actuar sobre los electrones son lo suficientemente importante para cambiar la distribución de probabilidad.

Vamos a usar una fuente menos brillante. Las ondas luminosas no perturbarán tanto a los electrones porque tienen menos energía. Lo primero que observamos es que los flashes reflejados por los electrones no se hacen más débiles. Lo único que ocurre es que algunas veces se oye el click en el detector sin haberse visto el flash. Es decir el electrón ha pasado sin

ser visto. Hemos observado que la luz actúa igual que los electrones, sabíamos que era onda, para, pero ahora encontramos que es corpúsculo. Siempre llega - o es reflejada - en corpúsculo que llamamos fotones. Si bajamos la intensidad de la luz no disminuirá el tamaño de los fotones, solamente la frecuencia con que son emitidos. Esto explica que cuando la fuente disminuye su intensidad, algunos electrones pasan sin ser vistos. Es decir no hay coincidencia entre el fotón y el electrón.

Si esto es verda, todo electrón que refleja luz es siempre igualmente perturbado. A pesar de todos vamos a realizar el experimento. Cuando oigamos un click en el detector lo contabilizaremos en una de estas tres columnas. En la 1 cuando el electrón sea visto cerca del hueco 1. En la 2 los electrones vistos cerca del hueco 2, y en la 3 los que no son vistos por ninguno de los dos. Cuando calculamos las probabilidades a partir de estos resultados encontramos: Para los vistos por el hueco 1 una distribución igual a P_1' ; para los que han pasado por 2 P_2' (o sea que los que son vistos por los huecos 1 y 2 tienen una distribución igual a P_{12}') y los que no son vistos por ninguno de los dos tienen una distribución ondulatoria igual a P_{12} de la figura 1.3. Si los electrones no son vistos hay interferencia!

Existirá alguna forma de ver los electrones sin perturbarlo. Como sabemos el momento asociado a un fotón es inmensamente proporcional a la longitud de onda ($P = h/\lambda$)

La desviación causada al electrón por el fotón cuando este es reflejado hacia nuestro ojo depende del momento que el fotón lleve asociado. Si deseamos que la perturbación sea mínima, no deberemos disminuir la intensidad, sino la frecuencia, (aumentando λ). Vamos a usar una luz más roja. También podríamos usar luz infraroja, u ondas de radio (como radar) y ver donde va el electrón por medio de algún equipo que pueda ver luz de estas λ mas largas.

Realizemos el experimento, repitiéndolo sucesivamente, cada vez con luces de mayor λ . Al principio obtenemos el mismo resultado. Entonces pasa una cosa terrible. Recordaremos que debido a la naturaleza ondulatoria de la luz, existe una limitación en la distancia entre dos puntos cercanos para que nuestro ojo los vea separados. Esta distancia es del orden de la longitud de onda de la luz. Cuando hagamos la longitud de onda mayor que la distancia entre los huecos veremos un gran flash y no distinguiremos por que hueco pasa el electrón. Y es precisamente con luz de este color, con la que perturbamos lo suficientemente poco a los electrones para que P_{12}' empiece a parecerse a P_{12} y tengamos efectos de interferencia. Y la curva P_{12} se

obtiene para luces de λ mucho mayor que la distancia entre los huecos.

Todo esto hizo que Heisenberg sugiriese que debería haber una ley en la naturaleza consistente en algunas limitaciones básicas en nuestras capacidades experimentales, no precisamente reconocidas. Propuso como principio general el principio de Incertidumbre, que pueda definirse en términos de nuestro experimento, como sigue: "Es imposible diseñar un aparato capaz de determinar por qué hueco pasa el electrón, que al mismo tiempo no perturbe lo suficiente a los electrones para destruir nuestro modelo de interferencia". Así debemos suponer que esto describe una característica básica de la naturaleza.

Toda la teoría de la mecánica cuántica depende de que este principio sea correcto ya que esta Teoría ha tenido éxito, este principio debe ser cierto.

Y qué pasa con la proposición A?. Bien, si disponemos del aparato, con luz, etc, es capaz de determinar cuando los electrones pasan por el hueco 1 o el 2, entonces se puede decir que pasan o bien a través de 1, o bien a través de 2. Pero si no intentamos encontrar el camino por el que van los electrones, cuando nuestro experimento no es perturbado, Entonces no se puede decir que un electrón pase a través de 1 o bien de 2. Si se dice esto o se empiezan a hacer deducciones a partir de aquí se caerá en el error.

1-7.- Primeros principios de mecánica cuántica

Vamos a resumir nuestras conclusiones. Definiremos un "experimento ideal" como uno en el que no hay influencias externas inciertas. Con mayor precisión diremos: "Un experimento ideal es aquel en el que todas las condiciones iniciales y finales del experimento están completamente especificadas". Lo que llamaremos "un suceso" es, en general, un conjunto específico de condiciones iniciales y finales (por ejemplo "un electrón deja el cañón, llega al detector y no pasa nada más).

RESUMEN

1) La probabilidad de un suceso en un experimento ideal viene dada por el cuadrado del valor absoluto de un número complejo Φ que se llama amplitud de probabilidad.

P = probabilidad

Φ = amplitud de probabilidad (1.6)

$$P = |\Phi|^2$$

2) Cuando un suceso puede ocurrir por varios caminos alternativos, la amplitud de probabilidad para el suceso es la suma de las amplitudes de probabilidad para cada camino consideradas separadamente. Existe interferencia.

$$\begin{aligned}\Phi &= \Phi_1 + \Phi_2 \\ P &= |\Phi_1 + \Phi_2|^2\end{aligned}\tag{1.7}$$

3) Si se realiza un experimento en el que somos capaces de determinar cuando se toma una u otra alternativa, la probabilidad de un suceso es la suma de las probabilidades para cada alternativa. Se pierde la interferencia.

$$P = P_1 + P_2\tag{1.8}$$

Debemos hacer incapié en una diferencia muy importante entre la mecánica clásica y la cuántica. Los físicos han renunciado a predecir exactamente que es lo que sucede en una circunstancia determinada, lo único que puede predecirse es la probabilidad de los diferentes sucesos.

Podríamos decir que esto es sólo por ahora -en el futuro quizás se descubra algún camino para evitar el principio de incertidumbre- pero hay fuentes sospechosas de que esto será para siempre. Este es el verdadero camino de la Naturaleza.

= . = . =

Ecuación de Onda de Schröndiger

Parte I : Introducción

2-1 Relaciones entre óptica y mecánica

El postulado de Sommerfeld conduce a la definición de una serie de números naturales característicos, para cada una de las partículas en movimiento (números cuánticos).

En 1.925 De Broglie introduce, a la par que la de la luz, la dualidad de la materia. Es decir considera que cada partícula con energía E , y cantidad de movimiento $P = mv$ lleva asociada una onda (ν, λ) tal que $E = h\nu$ $P = \frac{h}{\lambda}$ para tratar esta nueva dualidad De Broglie introduce dos definiciones diferentes de corpúsculo.

1º) El corpúsculo es un pequeño objeto, perfectamente localizado, que describe en el espacio y a lo largo del tiempo, una trayectoria, sobre la cual ocupa en cada instante una posición definida, estando animado de una velocidad determinada.

2º) El corpúsculo es una unidad Física caracterizada por ciertas constantes (masa, carga, etc) y susceptible de producir efectos localizados, donde interviene totalmente (y nunca por fracción) como por ejemplo el fotón en el efecto fotoeléctrico.

La segunda definición permite fácilmente asociar al concepto de partícula a la noción de onda: La intensidad de la onda asociada representa la probabilidad de que un fotón se manifieste.

El punto de partida de una de las formulaciones de la mecánica cuántica fué el "conjunto de puntos de De Broglie":

a) Existen analogías matemáticas (teoría de Hamilton-Jacobi) entre el principio de Fermat (desde donde se puede desarrollar toda la óptica geométrica), y el principio de Maupertuis (de donde puede deducirse la mecánica racional clásica).

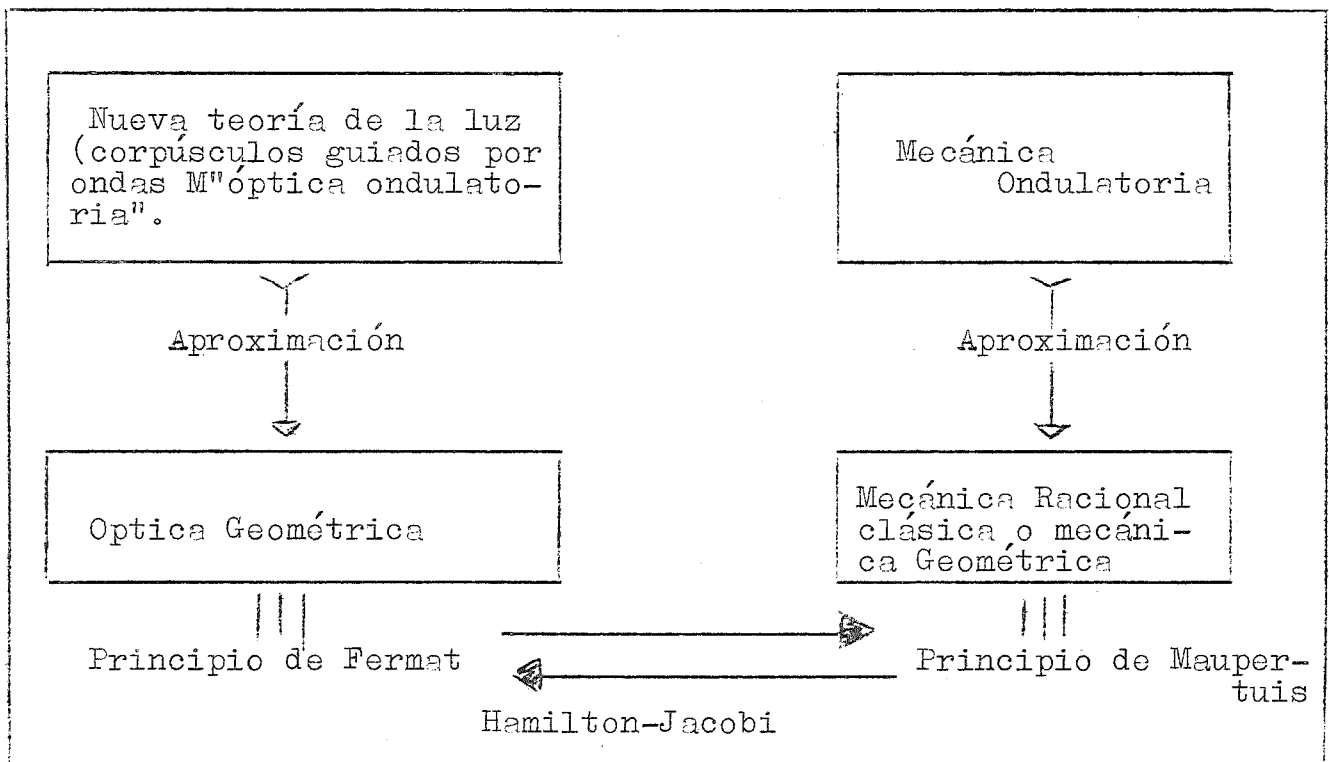
b) La óptica física, y por lo tanto el principio de Fermat son aproximaciones de la nueva teoría de la luz.

c) ¿Existirá una nueva mecánica de la cual el principio de Maupertuis y la mecánica clásica son aproximaciones?

La experiencia (vease capítulo I) ha confirmado la existencia de esta nueva mecánica (mecánica ondulatoria) más potente que la antigua.

Se puede resumir la situación en el cuadro siguiente.

Cuadro 2-1



Como es bien sabido el principio de Fermat dice que "el camino óptico entre dos puntos es mínimo" y el de Maupertuis es el principio de mínima acción (véase "Sandau-Mecánica" o "Goldstein-Mecánica clásica").

El fenómeno de interferencias, que tiene lugar cuando la luz se encuentra con objetos del mismo orden de magnitud que su λ , obliga a pasar de la Óptica geométrica a la óptica ondulatoria. El paso de la mecánica clásica a la mecánica ondulatoria es debido a un fenómeno parecido. En la siguiente sección veremos un ejemplo.

2-2 Ejemplos de aplicación.-

Para discernir cuando podrá aplicarse a un fenómeno el tratamiento aproximado de la mecánica clásica, tomaremos según indican los resultados experimentales, como figura de mérito Q/λ (Q coordenada generalizada, ver Sandau).

Supongamos un problema cuyas magnitudes son del orden de: cm., cm/seg., y gramo.

Según de Broglie $\lambda = h/p$, luego

$$\frac{Q}{\lambda} = \frac{Q p}{h} = \frac{Q m V}{h} = \frac{\text{cms. gr. cm/seg.}}{h} = \frac{\text{erg/seg}}{6,61 \times 10^{-27} \text{ erg/seg}} \gg 1$$

lo cual quiere decir por tanto que Q es muchísimo mayor que λ , por lo tanto la longitud de la onda asociada es extremadamente pequeña comparada con las dimensiones del fenómeno. Se podrá, por tanto aplicar la mecánica clásica.

Veamos ahora el caso de los dominios atómicos.

Por los resultados experimentales obtenidos sería más sencillo usar energías y Tiempos de transición o constantes de tiempo; así pues

$$E_T = 2 T = m V^2 = p V$$

(En el caso del oscilador armónico)

$$\tau = \frac{Q}{V} \quad \text{y nos quedará}$$

$$E \tau = m V^2 \frac{Q}{V} = p Q \quad \left\{ \begin{array}{l} \frac{E}{V} \approx 5 \text{ e V} \\ \nu \approx 10^{15} \end{array} \right.$$

$$\begin{aligned} \text{luego } \frac{p q}{h} &= \frac{E}{h} = \frac{F/\nu}{h} = \frac{c.V}{h\nu} = \frac{1,6 \times 10^{-19} \text{ V}}{6,61 \cdot 10^{-39} \text{ V}} = 2,4 \cdot 10^{14} \text{ V}/\nu = \\ &= \frac{2,9 \cdot 10^{19} \cdot \text{s}}{10^{15}} \approx 1,2. \end{aligned}$$

En este caso es evidente que la aplicación de la mecánica clásica nos conduciría a grandes errores.

Si utilizaráramos constantes termodinámicas tendríamos que usar las siguientes transformaciones.

$$K = 1,38 \times 10^{-16} \frac{\text{erg}}{\text{grado}} = 1,38 \cdot 10^{-23} \frac{\text{julios}}{\text{grado}}$$

y como $e V = K T$

Podríamos realizar una tabla en la que a cada voltaje correspondiera una temperatura

$$\text{a } T = 300^\circ \text{ K} \quad \text{Energía} \quad 0,025 \text{ V} = \frac{1}{40} \text{ V}$$

De este modo nos resulta que cuando

$$\frac{e V}{h \nu} = \frac{K T}{h \nu} \approx 1$$

El valor de T será

$$T = 500^\circ\text{K}$$

que es superior a la temperatura normal. Por encima de esta temperatura se puede utilizar la mecánica clásica y por debajo la ondulatoria. Esto ha tenido una brillante confirmación en la teoría de los calores específicos (ver KITTEL: Física del Estado sólido).

2-3 Velocidad de grupo y velocidad de partícula

En la teoría dual cada partícula lleva asociada una onda. Esta onda no se extenderá a todo el espacio, sino que será un paquete de ondas concentrado en el punto de máxima probabilidad de existencia de la partícula.

En un medio homogéneo este paquete de ondas se transmitirá siguiendo una trayectoria rectilínea. Pero si el índice de refracción varía de punto a punto, la trayectoria se curvará.

En óptica sucedería así. En mecánica ondulatoria será necesario encontrar alguna magnitud función de punto equivalente al índice de refracción óptico.

Para encontrarla supongamos en primer lugar que la frecuencia de la onda asociada permanece constante. En realidad esto lo podríamos justificar suponiendo un campo conservativo, en el cual la energía total se mantiene constante.

$$E = h \nu = C$$

luego será constante.

Al variar la velocidad de la partícula su longitud de onda λ también lo hará, ya que

$$\lambda = \frac{h}{p} = \frac{h}{m v}$$

De la fórmula de la energía total

$$E = T + V = \frac{1}{2} m v^2 + V = \frac{p^2}{2m} + V \quad \text{de donde } p = \sqrt{2m(E-V)}$$

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2m(E-V)}}$$

Al variar la energía potencial V cambiará el valor de λ , luego es este valor precisamente, la longitud de onda, la magnitud función de punto que buscábamos.

Esta idea física de analogía podrá llevarse mucho más lejos recurriendo a un desarrollo que partiendo del principio de Fermat nos conduzca al de mínima acción.

El Principio de Fermat se expresa

$$\int_{P_1}^{P_2} \frac{dl}{V} = 0$$

¿pero se le dice lo que físicamente representa este en-acc-on y el camino óptico?

o sea que el camino óptico es mínimo.

En esta fórmula V representa la velocidad de fase de la onda

$$V = \lambda \nu$$

luego

$$\oint_{P_1}^{P_2} \frac{\vec{dl}}{\lambda v} = \frac{1}{v} \oint_{P_1}^{P_2} \frac{\vec{dl}}{\lambda} = \frac{1}{v h} \oint_{P_1}^{P_2} p \vec{dl} = 0 \quad \text{lo que indica}$$

$$\oint_{P_1}^{P_2} p \vec{dl} = 0$$

Que es la expresión matemática del principio de mínima acción. Pero, a pesar de haber utilizado la velocidad de fase no nos dejemos engañar. Si utilizamos la velocidad de grupo llegaremos a resultados mucho más convenientes.

Definimos la velocidad de grupo del paquete de ondas como

$$V_g = \frac{d v}{d 1/\lambda}$$

Siendo su dirección normal al frente de ondas.

Asignemos un carácter vectorial a la longitud de onda, asignándole precisamente esta dirección

$$\frac{1}{\lambda} = (B_x, B_y, B_z) \text{ y por lo tanto } V_{gx} = \frac{\partial v}{\partial B_x};$$

$$V_{gy} = \frac{\partial v}{\partial B_y}; \quad V_{gz} = \frac{\partial v}{\partial B_z} \quad \text{ya que } v = \frac{E}{h} \text{ y } \frac{1}{\lambda} = \frac{p}{h}$$

tendremos que

$$B_x = \frac{p_x}{h}; \quad B_y = \frac{p_y}{h}; \quad B_z = \frac{p_z}{h} \quad \text{y} \quad V_{gx} = \frac{\partial E/h}{\partial \frac{p_x}{h}} = \frac{\partial E}{\partial p_x} = \frac{\partial H}{\partial p_x} \text{ y}$$

$$\text{análogamente } V_{gy} = \frac{\partial H}{\partial p_y}; \quad V_{gz} = \frac{\partial H}{\partial p_z}$$

Donde H es la función de Hamilton, equivalente a la energía total

$$H = \frac{1}{2} m v^2 + V = \frac{1}{2m} p^2 + V = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + V \text{ y por lo tanto}$$

$$V_{gx} = \frac{\partial H}{\partial p_x} = \frac{p_x}{m} = \frac{m v_x}{m} = v_x$$

luego hemos llegado a la conclusión de que la velocidad de grupo del paquete es la velocidad de la partícula considerada como tal. Luego si utilizamos esta representación, tendremos un resultado más coherente.

2-4 Principio de Incertidumbre

Corría el año 1.927 cuando Heisenberg (Zeits.F.Physik, 43,172 (1.927)) presentó el principio de incertidumbre. De acuerdo con este principio es imposible determinar exacta y simultáneamente los valores de los pares de variables físicas que determinan el comportamiento de un sistema atómico. Es decir aquellos pares de variables canó-

nicamente conjugado en el sentido de Hamilton (ver Landau: mecánica). La coordenada x de una partícula y el momento P_x ; La componente J_z del momento angular y la posición angular φ con la perpendicular al plano $x y$; La energía E de la partícula y el tiempo T en que se mide. Mas cuantitativamente, el principio postula además que el producto de los errores cometidos en las medidas nunca será menor que $\hbar/2$.

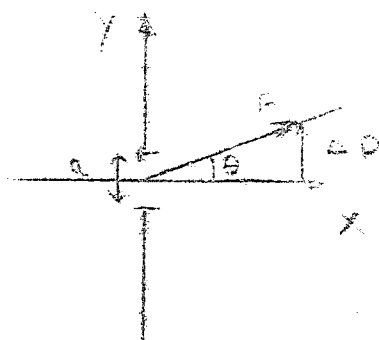
En resumen

$$\Delta x \cdot \Delta p_x \geq \hbar$$

$$\Delta \varphi \cdot \Delta J_z \geq \hbar$$

$$\Delta t \cdot \Delta E \geq \hbar$$

Supongamos el paquete de ondas asociado a la partícula que definimos anteriormente. Al pretender definir la posición de esta partícula, tendríamos que concentrar el paquete de ondas en un punto. Esto lo podríamos realizar experimentalmente por medio de diafragmas. Al hacer el orificio del diafragma cada vez mas pequeño para precisar la posición de la partícula se nos produce un fenómeno de difracción dentro de un caso de ángulo θ (fig. 1)



- 1 -

El momento o ímpetu puede oscilar entre un cierto número de valores y los límites de esta oscilación vienen dados según la figura por $\pm \Delta p$

Según leyes conocidas de la óptica, sobre difracción a través de una rendija, el ángulo θ vendrá dado por

$$\theta = \frac{\lambda}{a}$$

Entonces el error que cometemos en la medida de la posición de la partícula vendrá dado por la dimen-

sión de la rendija, o sea

$$\Delta y = a$$

De la figura vemos

$$\Delta p = p \sin \theta \approx p \theta$$

$$\Delta p \approx p \frac{\lambda}{a} = \frac{h}{\lambda} \frac{\lambda}{a} = \frac{h}{a}$$

de donde

$$\Delta y \Delta p = h$$

En el caso de máxima precisión vemos que no podemos medir con exactitud a la vez estos dos valores, y el resultado nos confirma el principio de Incertidumbre de Heisemberg anteriormente anunciado.

Vemos ahora lo que pasa con otras de las variables canónicas conjugadas que habíamos definido.

En el caso de tener un paquete de ondas, para precisar la posición de la partícula a la que habíamos asociado el paquete, tendríamos que concentrar este al máximo posible. Pero al concentrar el paquete disminuye el número de longitudes de onda que contiene y por lo tanto, aumenta el número de armónicos apreciables. Luego su aspecto se va desdibujando en frecuencia, en energía y por lo tanto en velocidad (ver fig. 2).



- 2 -

Este desdibujamiento en frecuencias se podría representar matemáticamente por la fórmula bien conocida.

$$\frac{\Delta \nu}{\nu} = \frac{1}{n^{\circ} \text{ de ondas}} = \frac{1}{\frac{c \Delta \tau}{\lambda}} = \frac{\lambda}{c \Delta \tau}$$

de donde

$$1 = \frac{\Delta \nu \Delta \tau c}{\lambda \nu} = \Delta \nu \Delta \tau \quad \text{pero}$$

$$\Delta \nu = \Delta \frac{E}{h} \quad \text{resultará.}$$

$$\frac{\Delta E}{h} \Delta \tau = 1 \quad \text{y evidentemente}$$

$$\Delta E \Delta \tau = h$$

De nuevo nos encontramos con esta limitación experimental, - que al parecer constituye una de las leyes fundamentales de la naturaleza.

2-5 De como, partiendo de la chapuza de la cuerda, llego Schödinger al establecimiento de su ecuación.

En la vieja mecánica cuántica, la cuantificación se obtenia a partir de la integral de fase de Somnerfeld

$$\oint p \, dq = n h$$

El número de longitudes de onda en una determinada distancia será

$$\int \frac{dq}{\lambda} \quad \text{ya que } p = h / \lambda \quad \text{luego}$$

$$\int \frac{dq}{\lambda} = \int \frac{p \, dq}{h} = n^{\circ} \text{ de longitudes de onda}$$

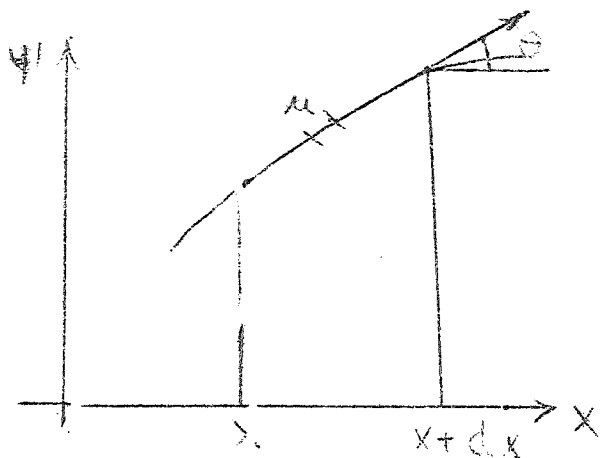
Si exigimos que estas ondas sean estacionarias el número de semilargitudes de onda tiene que ser entera, o dicho matemáticamente

$$\int p \frac{dq}{h} = n \quad \oint p \, dq = n h$$

!oh, maravilla, la integral de fase de Somnerfeld resulta ser idéntica a la exigencia de los procesos de ondas estacionarias!.

Si esto es así, dijo Schödinger, el tratamiento matemático utilizado para estas ondas, debe ser válido para dichos procesos cuánticos.

Un problema clásico de ondas estacionarias es el de los modos de vibración de la cuerda tensa, cuya resolución se ataca como sigue



Llamando Ψ a la separación de un punto de suposición de reposo después de haber aplicado una - Fuerza, y haciendo uso de la ecuación fundamental de la mecánica racional

$$\vec{F} = m \vec{a} \quad \text{tenemos}$$

$$\lim_{\Delta x \rightarrow 0} \frac{T \left(\frac{\partial \Psi}{\partial x} \right)_{x+\Delta x} - T \left(\frac{\partial \Psi}{\partial x} \right)_x}{\Delta x} =$$

$$= T \left(\frac{\partial^2 \psi}{\partial x^2} \right) \quad F = T \left(\frac{\partial^2 \psi}{\partial x^2} \right) dx$$

y llamando μ la masa por unidad de longitud

$$F = m.a = \mu dx \frac{\partial^2 \psi}{\partial \tau^2}$$

luego la ecuación diferencial de la cuerda vibrante es

$$T \left(\frac{\partial^2 \psi}{\partial x^2} \right) = \mu \frac{\partial^2 \psi}{\partial \tau^2}$$

Son soluciones que nos interesarán serán las periódicas con respecto al tiempo, del tipo

$$\psi(x, t) = \mu(x) \sin \omega \tau$$

Sustituyendo en la ecuación tendremos.

$$\left. \begin{aligned} \frac{\partial^2 \psi}{\partial \tau^2} &= -\omega^2 \psi = -\omega^2 \mu(x) \text{ en } \omega \tau \\ \frac{\partial^2 \psi}{\partial x^2} &= \frac{d^2 \mu}{dx^2} \sin \omega \tau \end{aligned} \right\} \text{ luego}$$

$$T \frac{d^2 \mu}{dx^2} = -\mu \omega^2 \mu(x) \quad \text{y} \quad \left\{ \frac{d^2 \mu(x)}{dx^2} + \frac{\mu}{T} \omega^2 \mu(x) = 0 \right\}$$

Al resolver esta ecuación obtenemos soluciones del tipo de propagación de ondas

$$\mu = \sin 2\pi \frac{x}{\lambda} \quad \text{o sea}$$

$$\frac{2\pi}{\lambda} = \omega \sqrt{\mu/T}$$

de donde, finalmente y restituyendo las derivadas parciales obtenemos.

$$\left\{ \frac{\partial^2 \mu}{\partial x^2} + \left(\frac{2\pi}{\lambda} \right)^2 \mu = 0 \right\}$$

Schrödinger, haciendo uso de la solución de de Broglie

$$E = h \nu \quad ; \quad P = h/\lambda$$

realizó las siguientes transformaciones

$$E = \frac{1}{2} m v^2 + V = \frac{p^2}{2m} + V \quad \text{de aquí}$$

$$P = \sqrt{2m(E-V)} \quad \frac{1}{\lambda} = \frac{\sqrt{2m(E-V)}}{h} \quad \text{y sustituyendo en la ecuación de onda}$$

$$\frac{\partial^2 \mu}{\partial x^2} + \frac{2\pi^2 m}{h^2} (E-V) \mu = 0$$

Esta ecuación nos da la longitud de la onda en cada punto x en un instante considerado ($t = \text{cte}$) puesto que en caso más general sería función de punto y del tiempo.

Podríamos generalizar esta ecuación a tres dimensiones

$$\Delta \psi + \frac{2m}{\hbar^2} (E - V) \psi = 0$$

Esta es la expresión más sencilla de esta ecuación que se conoce con el nombre de primera ecuación de Schrödinger.

2-6 Planteamiento de la segunda ecuación de Schrödinger

En mecánica clásica se expresa la energía total de una partícula por una función llamada Hamiltoniano H (ver Goldstein : mecánica-clásica)

$$H = \frac{1}{2m} (P_x^2 + P_y^2 + P_z^2) + V(x, y, z)$$

Schrödinger observó que si se introducen una correspondencia entre las componentes de la cantidad de movimiento y ciertos operadores diferenciales, expresada de este modo:

$$P_x \rightarrow \frac{\hbar}{2\pi i} \frac{\partial}{\partial x}$$

$$P_y \rightarrow \frac{\hbar}{2\pi i} \frac{\partial}{\partial y}$$

$$P_z \rightarrow \frac{\hbar}{2\pi i} \frac{\partial}{\partial z}$$

la función de Hamilton podría escribirse $H \rightarrow \frac{-\hbar^2}{8\pi^2 m} \Delta + V$

y entonces, aplicando la nueva igualdad mecánica $H = E$ a la función de onda ψ

$$H\psi = E\psi$$

Nos resulta la ecuación de Schrödinger $\frac{\hbar^2}{8\pi^2 m} \Delta \psi + (E - V) \psi = 0$

Evidentemente esta representación excepcional de la ecuación de Schrödinger, adolece del defecto de no poderse aplicar en el caso de que deseemos restituir la variable temporal en la ecuación de onda.

Observemos que lo que hemos hecho es definir una correspondencia entre el momento \vec{P} , o mejor dicho, sus componentes P_x , P_y , P_z y los operadores derivada con respecto a sus variables canónicas conjugadas x, y, z .

Si consideramos una concepción espacio-temporal, plenamente relativista, los vectores que representan un punto en el espacio de cuatro dimensiones incluye tres componentes espaciales (x, y, z) y una componente temporal,

Como se demuestra en mecánica las derivadas de la acción

S , $\frac{\partial S}{\partial x}$, $\frac{\partial S}{\partial y}$, $\frac{\partial S}{\partial z}$, son las tres componentes del impulso de la partícula y la derivada $-\frac{\partial S}{\partial t}$ es la energía de la misma.

Por lo tanto podríamos definir en mecánica relativista, otros cuadrivectores cuyas componentes espaciales coinciden con las del impulso \vec{P} y cuya componente temporal viene representada por $-E$. Esto coincide con el hecho de ser la energía variable canónica conjugada del tiempo.

Entonces, a las correspondencias expresadas anteriormente podríamos añadir una cuarta

$$E \longrightarrow - \frac{h}{2\pi j} \frac{\partial}{\partial \tau}$$

y la ecuación $H \psi = E \psi$ nos queda de la forma.

$$-\frac{h^2}{8\pi^2 m} \Delta \psi + V \psi = - \frac{h}{2\pi j} \frac{\partial \psi}{\partial \tau}$$

que es la segunda ecuación de SchRödinger, donde la función de onda ψ es dependiente del tiempo.

A. FORMALISMO GENERAL DE LA TEORIA CUANTICA

No daremos aquí mas que unas breves nociones del formalismo general empleado en el desarrollo de la Teoría cuántica, con exposición de los principales teoremas en los que se basa, sin profundizar en ellos ni obtener conclusiones particulares. Nos basaremos en la exposición hecha por Dirac, con su rotación y los puntos fundamentales que trata. Para la división en conceptos matemáticos necesarios y conceptos físicos, seguiremos el planteamiento de Meniah.

A-1 INTRODUCCION

En física Clásica, el estado dinámico de un sistema dado viene definido en cada instante cuando se conocen en él los valores que toma el conjunto de las variables dinámicas. Estos pueden ser todos simultáneamente determinados con una posición infinita. El objeto de la Teoría Clásica es enumerar estas variables dinámicas y después descubrir sus ecuaciones de evolución.

En Teoría Cuántica, la relación entre estados dinámicos y variables dinámicas es mucho menos directo. En la operación de medida de una variable dinámica dada, el estado dinámico del sistema sobre el que se efectúa la medida se encuentra en general modificada por la intervención del aparato de medida. Esta modificación, corrientemente despreciada en Física Clásica, deja de serlo aquí en la escala microscópica; aparece como una perturbación imprescindible e incontrolable del sistema y establece un límite a la precisión con la cual pueden ser medidas simultáneamente las dichas variables dinámicas. Esto sirve a Dirac para definir los conceptos de grande y pequeño de una manera absoluta. Según este, a fin de dar un significado absoluto al tamaño, tal como requiere una teoría de la estructura última de la materia, hemos de asumir que hay un límite para la precisión de nuestro poder de observación y la pequeñez de la perturbación que la acompaña: límite que es inherente a la naturaleza de las cosas y no puede ser nunca sobrepasado por mejoras en la Técnica o del cuidado del observador. Si el objeto bajo observación es tal que la perturbación límite incluíble es despreciable, entonces el objeto es grande en el sentido absoluto y podremos aplicarle la Mecánica Clásica.

Si, por otra parte, la perturbación límite no es despreciable, entonces el objeto es pequeño en el sentido absoluto y requerirá los procedimientos de la mecánica cuántica.

Se abandona con ello los postulados de la Mecánica Clásica según los cuales las variables dinámicas toman en un cierto instante valores perfectamente determinados, para tomar otros según las cuales solo fijaremos para cada una de estas una distribución estadística de valores, dando la ley de probabilidad de los resultados de la medida en la eventualidad de que una medida tal sea efectuada.

Habrán ahora unas ciertas variables dinámicas que pueden coexistir con otras, esto es, que podemos efectuar medidas de ambas simultáneamente. Diremos de ellas que son compatibles. Pero habrá otras que la medida que hagamos a una afectará a la otra de forma que no podamos efectuar medidas precisas sobre ella. En este caso se trataría de variables incompatibles. Veremos que condición habrá de verificar. Se admite que se puede asignar a cada variable dinámica del sistema, un cierto número de otras y formar con ellas un sistema completo de variables compatibles; por definición, todas las variables de un sistema tal son compatibles dos a dos. La determinación precisa de las variables de un conjunto completo constituye la mas grande suma de información que se puede obtener sobre el estado dinámico de un sistema cuántico. En consecuencia, el estado dinámico de un sistema cuántico queda definido, no como en teoría clásica conociendo de forma precisa, todas las variables relacionadas con el sistema, sino solo aquellas que figuran en uno de los diversos conjuntos completos que se pueden formar.

Se supone, en primer lugar que los estados dinámicos de un sistema cuántico son linealmente superponibles. La naturaleza no clásica del proceso de superposición se comprende mas claramente considerando la superposición de dos estados A y B tales que existe una observación que hecha cuando el sistema está en A, conduce al resultado a y cuando hecha en B, da el b. ¿ Cual será ahora el resultado de la observación cuando se hace estando el sistema en el estado compuesto?. La respuesta es que a veces estará en a y a veces en b, de acuerdo a una ley de pro

babilidad que dependerá de los pesos relativos de A y B en la superposición. Nunca será diferente de ambos, a y b . El caracter intermedio del estado formado por superposición expresa por ello la probabilidad de un resultado particular en una observación o sea intermedia entre las correspondientes probabilidades para los estados originales, pero no que el resultado mismo será intermedio entre los primitivos para los estados originales.

El proceso de superposición es pues de tal índole que posee caracter aditivo: implica que puedan sumarse estados de forma que den nuevos estados. Los estados habrán de conectarse con cantidades matemáticas que puedan sumarse dando otras cantidades de la misma naturaleza. Las mas obvias son los vectores. Pero los vectores ordinarios, existentes en un espacio de número finito de dimensiones, no son suficientemente generales para la mayor parte de los sistemas dinámicos en M.Q. Habríamos de generalizar a espacios de número infinito de dimensiones y el tratamiento matemático pasaría a ser complicado por cuestiones de convergencia.

Será pues conveniente tratar con vectores de unas ciertas características, para conectarlos convenientemente con los estados de un sistema. A estos vectores los denominaremos vectores K E T o mas concretos K E T S, designándolos con el símbolo $| \rangle$ y si se desea especificar uno, como A . Estos Kets podrán ser multiplicados por números complejos y sumados con otros Kets, para dar otros vectores Kets, esto es

$$C_1 | A \rangle + C_2 | B \rangle = | R \rangle$$

que no es mas que la expresión del principio de superposición. Estados que resultan de la superposición de otros se dicen dependen de ellos. Mas en general, diremos que un estado es dependiente de un conjunto de estados, finito o infinito en número, cuando el correspondiente Ket depende de los Kets de los correspondientes estados. Un conjunto de estados se denomina independiente si no depende de otros.

Como consecuencia de esta formulación matemática hemos de hacer la nueva suposición de que la superposición de un estado consigo mismo no puede dar un nuevo estado, sino el mismo de nuevo. Si el estado original es el $| A \rangle$, cuando superpuesto consigo mismo dará

$$C_1 |A\rangle + C_2 |A\rangle = (C_1 + C_2) |A\rangle$$

como nuestro requerimiento es que el nuevo estado $(C_1 + C_2) |A\rangle$ sea el mismo que el original $|A\rangle$, y $C_1 + C_2$ es un número complejo arbitrario hemos de inferir que si un Ket cualquiera correspondiente a un cierto estado es multiplicado por cualquier número complejo, no cero, el Ket resultante corresponderá al mismo estado. Por ello un estado está especificado por una dirección de un vector Ket y cualquier longitud que queramos asignarles es indiferente a efectos de descripción del estado. Todos los estados de un sistema dinámico están en correspondencia binnívoca con todas las posibles direcciones para un vector Ket.

El espacio que nos interesa pues es uno en el que no existe límite en el número de vectores linealmente independientes y que como consecuencia tendrá dimensión infinita. Esta condición, y otras que también requeriremos son satisfechas por el espacio de Hilbert.

A- 2 ESPACIO DUAL

Como ya se sabe, cuando tenemos un conjunto de vectores en cualquier teoría matemática, podemos construir un segundo conjunto de nuevos vectores relacionado con el anterior y que recibe el nombre de "espacio dual".

Llamaremos a los vectores de este nuevo espacio como vectores BRA, o simplemente bras, representándolos por $\langle B|$, y si queremos concretar o uno $\langle B|$.

El producto escalar del bra $\langle B|$ por el Ket $|A\rangle$ será escrito como $\langle A|B\rangle$. Vemos así que un bra-ket completo denota un número y cualquier braket incompleto denota un vector, bra o ket según la parte que hayamos suprimido.

La condición de que el producto escalar de $\langle B|$ y $|A\rangle$ es función lineal de $|A\rangle$ puede expresarse simbólicamente por

$$\langle B| \{ |A\rangle + |A'\rangle \} = \langle B|A\rangle + \langle B|A'\rangle$$

$$\langle B| \{ C |A\rangle \} = C \langle B|A\rangle \quad c = \text{cualquier número}$$

Un vector bra se considera completamente definido cuando se da un producto escalar en cada vector ket, o mas particularmente, con cada vector ket de la base que constituye el espacio vectorial considerado.

Entre estos espacios asumimos se puede efectuar una correspondencia binnívoca de forma que al bra correspondiente a $|A\rangle + |A'\rangle$ le corresponde la suma de los bras correspondientes a $|A\rangle$ y a $|A'\rangle$ y el correspondiente al $C|A\rangle$ es C veces el bra correspondiente al $|A\rangle$.

Esta correspondencia es antilineal, esto es, el bra conjugado del Ket.

$$|A\rangle = a_1 |1\rangle + a_2 |2\rangle \quad \text{es el}$$

$$\langle A| = a_1^* \langle 1| + a_2^* \langle 2|$$

De acuerdo con esta correspondencia, cualquier estado de nuestro sistema dinámica puede ser especificado lo mismo por la dirección de un vector bra que por la de un ket.

Dado cualquier conjunto de dos vectores $|A\rangle$ y $|B\rangle$, podemos construir de ellos el número $\langle B|A\rangle$ tomando el producto escalar del primero con el conjugado imaginario del segundo. Este número depende linealmente de $|A\rangle$ y antilinealmente de $|B\rangle$; la dependencia antilineal significa que el número formado $|B\rangle + |B'\rangle$ es la suma de los formados por $|B\rangle$ y por $|B'\rangle$, y el número formado por $C|B\rangle$ es c veces el formado por $|B\rangle$.

Un segundo camino de construir un número con esos dos vectores es tomar el producto escalar de $|B\rangle$ con el conjugado imaginario de $|A\rangle$ y tomando el complejo conjugado de este producto escalar. Supondremos que estos dos números son iguales, esto es

$$\langle B|A\rangle = \langle A|B\rangle^*$$

Poniendo $|B\rangle = |A\rangle$, encontramos que el número $\langle A|A\rangle$ debe ser real. Se hará además la suposición $\langle A|A\rangle > 0$

Diremos que dos vectores son ortogonales si su producto escalar es cero.

El módulo de un vector $|A\rangle$ se define como la raíz cuadrada del número positivo $\langle A|A\rangle$

De estas propiedades resulta la desigualdad de Schwarz:

$$|\langle u|v\rangle| \leq \sqrt{\langle u|u\rangle \langle v|v\rangle}$$

Los axiomas que preceden deben ser completados por la hipótesis de que el espacio de los vectores kets E es completo y separable. Con ello nos resulta que el espacio formado, no es otro que el espacio de Hilbert.

A- 3 OPERADORES LINEALES

Supongamos que a cada ket $|A\rangle$ del espacio vectorial corresponde un cierto ket $|B\rangle$: se dice que $|B\rangle$ resulta de la acción de un cierto operador sobre el ket $|A\rangle$. Si además, esta correspondencia es lineal, el operador así definido es un cierto operador lineal α . Se escribe

$$|B\rangle = \alpha |A\rangle$$

Daremos ahora unas ciertas propiedades de estos operadores.

A.3.1. Para que un operador sea nulo, es necesario y suficiente que $\forall |A\rangle$ se tenga $\langle A | \alpha | A \rangle = 0$

A.3.2. Para que dos operadores, α y β sean iguales, es necesario y suficiente que $\forall |A\rangle$ se tenga $\langle A | \alpha | A \rangle = \langle A | \beta | A \rangle$

A.3.3. Las principales operaciones de algebra sobre los operadores son:

a) Multiplicación de un operador α por una constante c

$$(c\alpha)|A\rangle \equiv c(\alpha|A\rangle)$$

b) Suma $\Sigma = \alpha + \beta$ de dos operadores α y β :

$$\Sigma|A\rangle \equiv \alpha|A\rangle + \beta|A\rangle$$

c) Producto $\Pi = \alpha\beta$ de un operador β por otro α

$$\Pi|A\rangle \equiv \alpha\beta|A\rangle \equiv \alpha(\beta|A\rangle)$$

que tiene la propiedad fundamental de que no es conmutativo, diferencia esencial con el algebra de las operadores lineales y el algebra ordinaria. El producto $\alpha\beta$ no es necesariamente igual al producto $\beta\alpha$. La diferencia $\alpha\beta - \beta\alpha$ de estas dos cantidades se designa con el nombre de conmutador de α y β , representándole por el símbolo (α, β) :

$$(\alpha, \beta) \equiv \alpha\beta - \beta\alpha$$

Si esta diferencia es nula, se dice que los dos operadores conmutan.

Las principales reglas que gobiernan el algebra de los conmutadores, y que serán aplicadas luego, son

$$c.1) \quad (\alpha, \beta) = -(\beta, \alpha)$$

$$c.2) \quad (\alpha, \beta + \gamma) = (\alpha, \beta) + (\alpha, \gamma)$$

$$c.3) \quad (\alpha, \beta \gamma) = (\alpha, \beta) \gamma + \beta (\alpha, \gamma)$$

$$c.4) \quad (\alpha, (\beta, \gamma)) + (\beta, (\gamma, \alpha)) + (\gamma, (\alpha, \beta)) = 0$$

Por aplicación repetida de la regla c.3, se obtiene

$$(\alpha, \beta^n) = \sum_{s=0}^{n-1} \beta^s (\alpha, \beta) \beta^{n-s-1}$$

---.---.---.---.---

Mas antes de seguir adelante, recapitulemos y obtengamos nuevas consecuencias haciendo nuevas suposiciones.

Hemos completado un esquema algebraico relacionando tres clases de cantidades: vectores bra, vectores ket y operadores lineales. Supusimos que los vectores bra o los ket, o mejor, sus direcciones, corresponden a los estados de un sistema dinámico en un momento determinado. Ahora haremos la suposición de que "los operadores lineales se corresponden con las variables dinámicas del sistema en ese momento". Entendiendo por variables dinámicas cantidades tales como las coordenadas y las componentes de la velocidad, cantidad de movimiento y momento angular de las partículas, y funciones de estas cantidades.

---.---.---.---.---

Si suponemos ahora que la correspondencia definida antes entre $|A\rangle$ y $|B\rangle$ es binnívoca, define dos operadores lineales α y β :

$$|B\rangle = \alpha |A\rangle \quad |A\rangle = \beta |B\rangle$$

estos operadores, son por definición inversos uno del otro, y verifican que $\alpha\beta = \beta\alpha = 1$. El inverso de un operador no existe siempre. Cuando existe se le designa por α^{-1} . Con lo anterior puede demostrarse que el producto de dos operadores α y β que poseen inversos, tiene inverso y es de la forma: $(\alpha\beta)^{-1} = \beta^{-1} \alpha^{-1}$

A.4 OPERADORES ADJUNTOS Y RELACIONES DE CONUGACION.-

Sea α un operador lineal. Sea $|B\rangle$ el ket conjugado del bra $\langle A|\alpha$. $|B\rangle$ depende antilinealmente del bra $\langle A|$, es pues una función lineal de $|A\rangle$. Esta correspondencia lineal define un operador lineal al que se da el nombre de operador "conjugado hermítico" de α o operador adjunto de α , designándole por α^\dagger :

$$|B\rangle = \alpha^\dagger |A\rangle$$

d

Ya que $\alpha^\dagger |A\rangle$ es el ket conjugado del bra $\langle A|\alpha$, el producto escalar de este ket por un bra arbitrariamente dado $\langle C|$ es complejo conjugado del de $|C\rangle$ por $\langle A|\alpha$, llegándose así a

$$\langle C|\alpha^\dagger |A\rangle = \overline{\langle A|\alpha |C\rangle}$$

De manera análoga se obtienen las relaciones fundamentales siguientes

$$\begin{aligned} (c\alpha)^\dagger &= c^* \alpha^\dagger \\ (\alpha + \beta)^\dagger &= \alpha^\dagger + \beta^\dagger \\ (\alpha\beta)^\dagger &= \beta^\dagger \alpha^\dagger \\ (\alpha^\dagger)^\dagger &= \alpha \end{aligned}$$

A.5 OPERADORES HERMITICOS.-

Por definición, un operador lineal \mathcal{E} es hermitico si es su propio adjunto:

$$\mathcal{E} = \mathcal{E}^\dagger$$

Es antihermitico si es opuesto a su propio adjunto.

$$\mathcal{E} = -\mathcal{E}^\dagger$$

A.6 VALORES PROPIOS Y VECTORES PROPIOS (EIGENVALORES Y EIGENVECTORES.-

RES.-

Vamos ahora a efectuar un nuevo desarrollo de la teoria de los operadores lineales, con el estudio de la ecuación.

$$\alpha |P\rangle = a |P\rangle$$

donde α es un operador lineal y a un número complejo. Esta ecuación se presenta normalmente de forma que α se conoce y el número a y $|P\rangle$ son desconocidos. El problema consiste en encontrar a estos, dejando aparte la solución trivial $|P\rangle = 0$.

Esta ecuación significa que el operador lineal α aplicado al ket $|P\rangle$ le multiplica por un feutor numérico sin cambiar su dirección. Este mismo α , es lógico aplicado a otros kets los cambiará en dirección y módulo.

Si se cumple la ecuación anterior, llamaremos a a un valor propio (eigenvalor) del operador lineal α , o de la correspondiente variable dinámica, y a $|P\rangle$ el vector ket propio de ese operador lineal.

Diremos también que el ket propio $|P\rangle$ pertenece al valor propio a .

Si es $|P\rangle$ un ket propio de \mathcal{A} , todo múltiplo $c|P\rangle$ de este, será también ket propio de \mathcal{A} relativo al mismo valor propio; si existen varios kets propios linealmente independientes relativos al mismo valor propio a , toda combinación lineal de ellos es ket propio de \mathcal{A} relativo a este valor propio. En otros términos, el conjunto de kets propios de

relativos a un valor propio dado a forma un subespacio vectorial relativo a ese valor propio a . Si ese subespacio no tiene mas que una dimensión, se dice que el valor propio es no degenerado. Si no, que hay degeneración, viniendo dada la degeneración por el número de dimensiones de este subespacio.

El único caso de interés práctico es cuando \mathcal{A} es hermitico. Entonces

- i) los dos espectros de valores propios del espacio de los bra y de los kets son idénticos.
- ii) todos los valores propios son reales.
- iii) el subespacio de bras propios relativos a un valor propio dado es el dual del subespacio de los kets propios relativos al mismo valor propio.

La demostración de ii), bastante trivial es:

Si $\mathcal{A} = \mathcal{A}^\dagger$ y $\mathcal{A}|A\rangle = a|A\rangle$ multiplicando por $\langle A|$

$$\langle A|\mathcal{A}|A\rangle = a\langle A|A\rangle$$

como $\langle A|\mathcal{A}|A\rangle^* = \langle A|\mathcal{A}^\dagger|A\rangle = \langle A|\mathcal{A}|A\rangle$

habrá de ser $\langle A|\mathcal{A}|A\rangle$ real y como ya lo era $\langle A|A\rangle$, será también real.

Una propiedad importante de los vectores propios de un operador hermitico es la de ortogonalidad de los vectores propios relativos a dos valores propios distintos. En efecto, si es

$$\mathcal{A}|A\rangle = a|A\rangle \quad \mathcal{A}|B\rangle = b|B\rangle \quad \text{ó} \quad \langle B|\mathcal{A} = b\langle B|$$

multiplicando la prueba escalarmente por $\langle B|$ y la segunda por $|A\rangle$ que da $0 = (a-b)\langle B|A\rangle$ y si $a \neq b$ $\langle B|A\rangle = 0$ c.q.d

A.7 OBSERVABLES.-

Si el conjunto de los vectores propios de un operador hermitico subtiende todo el espacio, esto es, si todo vector de norma finita puede ser desarrollado en serie de estos vectores, se dicen que forman un sistema completo y que el operador hermitico es un "observable".

Los observables son los únicos operadores hermiticos del espacio E susceptibles de tener una interpretación física.

1.8 OPERADORES QUE CONMUTAN CON UN OBSERVABLE, OBSERVABLES QUE CONMUTAN.-

Sea \mathcal{E} un observable y \mathcal{E}' un valor propio del mismo correspondiente al vector propio $|\mathcal{E}'\rangle$, esto es

$$\mathcal{E} |\mathcal{E}'\rangle = \mathcal{E}' |\mathcal{E}'\rangle$$

Si se verifica $(\mathcal{E}, X) = 0$, se tendrá evidentemente

$$0 = (\mathcal{E} X - X \mathcal{E}) |\mathcal{E}'\rangle = \mathcal{E} X |\mathcal{E}'\rangle - X \mathcal{E} |\mathcal{E}'\rangle = \mathcal{E} X |\mathcal{E}'\rangle - X \mathcal{E}' |\mathcal{E}'\rangle = \\ = \mathcal{E} X |\mathcal{E}'\rangle - \mathcal{E}' X |\mathcal{E}'\rangle = (\mathcal{E} - \mathcal{E}') X |\mathcal{E}'\rangle \Rightarrow \mathcal{E} (X |\mathcal{E}'\rangle) = \mathcal{E}' (X |\mathcal{E}'\rangle)$$

esto es $X |\mathcal{E}'\rangle$ es vector propio de \mathcal{E} relativo al mismo valor propio.

Esto se aplica, en particular, al caso de observables que conmutan. Los resultados que pueden obtenerse y a los que no llegaremos aquí en su obtención detallada son:

Se llama sistema de base de un observable dado todo el sistema ortonormal completo de vectores propios de este observable. Con esta definición, se tiene el siguiente teorema Th: Si dos observables \mathcal{E}_1 y \mathcal{E}_2 conmutan, poseen al menos un sistema de base común y, reciprocamente, si dos observables \mathcal{E}_1 y \mathcal{E}_2 poseen un sistema de base común, conmutan.

1.9 ESTADOS DINAMICOS Y MAGNITUDES FISICAS.-

1.9.1. DEFINICION DE PROBABILIDADES. POSTULADOS RELATIVOS A LA MEDIDA.-

A cada estado dinámico corresponde una cierta distribución estadística de valores para cada una de las variables dinámicas pertenecientes al sistema. Para definir estas probabilidades, se establece el postulado fundamental siguiente:

El valor medio de una función $F(A)$ de una magnitud física dado es

$$\langle F(A) \rangle = \langle X | F(A) | X \rangle \quad (\text{A.q.1.a.})$$

expresión en la cual el ket $|X\rangle$ representa el estado dinámico del sistema y el observable A la magnitud física en cuestión.

En particular la función característica $F(\mathcal{E})$ de la distribución estadística de A es el valor medio de la función $e^{i \mathcal{E} A}$.

¿Dónde aparece esto?

131

ESTEREOELECTRONICA

$$\frac{d}{dq} (q|\psi\rangle) = q \frac{d}{dq} |\psi\rangle + |\psi\rangle \quad \text{y} \quad \frac{d}{dq} q|\psi\rangle - q \frac{d}{dq} |\psi\rangle = |\psi\rangle$$

$$\left(\frac{d}{dq} \cdot q - q \frac{d}{dq} \right) |\psi\rangle = |\psi\rangle \quad \text{y como ha de ser } \forall |\psi\rangle \text{ verificará}$$

$$\frac{d}{dq} q - q \frac{d}{dq} = 1 \quad \text{ó} \quad \left(\frac{d}{dq} \cdot q \right) = 1 \quad (\text{A.12.b})$$

relación que comparada con la (A.12.a) queda que $-i\hbar \frac{d}{dq}$ satisface la misma relación de conmutación que q con p .

Por ello podemos tomar

$$\boxed{p \rightarrow -i\hbar \frac{d}{dq}} \quad (\text{A.12.c})$$

sin causar ninguna perturbación en el esquema presentado. Y esto será lo que hagamos en lo sucesivo.

Como aplicación veamos el valor del conmutador de p con una función de q . Se cumplirá evidentemente que

$$\frac{\partial}{\partial q} F|\psi\rangle = F \frac{\partial}{\partial q} |\psi\rangle + \frac{\partial F}{\partial q} |\psi\rangle \quad \text{y como es } \forall |\psi\rangle \text{ será}$$

$$\frac{\partial}{\partial q} F - F \frac{\partial}{\partial q} = \frac{\partial F}{\partial q} \quad \text{y con ayuda de la expresión anterior}$$

$$\text{queda} \quad p F - F p = -i\hbar \frac{\partial F}{\partial q} \quad \text{o lo que es igual}$$

$$\boxed{(F, q) = i\hbar \frac{\partial F}{\partial q}} \quad (\text{A.12.d})$$

llegándose igual a

$$\boxed{(p, F(q)) = i\hbar \frac{\partial F}{\partial q}} \quad (\text{A.12.e})$$

B. EL OSCILADOR ARMONICO

B.1 INTRUDUCCION

Consideremos una partícula ejecutando pequeñas oscilaciones en una dimensión (lo que se denomina un oscilador lineal). La energía potencial de una partícula en esas condiciones viene dada en Mecánica Clásica por $\frac{1}{2} m \omega^2 x^2$ donde ω es la frecuencia (circular) característica de las oscilaciones. De acuerdo con esto, el Hamiltoniano del oscilador es

$$\textcircled{H} = \frac{1}{2} \frac{p^2}{m} + \frac{1}{2} m \omega^2 x^2$$

El estudio cuántico de un tal sistema presenta la ventaja de que puede resolverse sigurosamente su ecuación de Schödinger y además de que interviene en todos los problemas en los que entran en juego oscilaciones cuantificadas: se le encuentra en Electrodinámica Cuántica, en Teoría Cuántica de Campos, en la teoría de las vibraciones moleculares y cristalinas. Además, los problemas que se presentan en su estudio son una excelente ilustración de todos los principios que hemos visto hasta aquí.

B.2 EL PROBLEMA DE LOS VALORES PROPIOS

A fin de evitar constantes inútiles en los cálculos, pondremos

$$H = \hbar \omega \mathcal{H}$$

$$q = \left(\frac{\hbar}{m \omega} \right)^{\frac{1}{2}} Q \quad p = (m \hbar \omega)^{\frac{1}{2}} \mathcal{P}$$

con lo que el Hamiltoniano anterior quedará

$$H = \frac{1}{2} (\mathcal{P}^2 + Q^2) \quad (\text{B.2.a})$$

nuestro problema será calcular los valores propios y los vectores propios de este operador.

Como puede verse facilmente los operadores hermíticos satisfacen las relaciones de conmutación:

$$(Q, \mathcal{P}) = i \quad (\text{B.2.b})$$

Para resolver el problema, tomaremos la representación de Schödinger, en el que la ecuación de Schödinger quedará

$$\frac{1}{2} \left(-\frac{d^2}{dQ^2} + Q^2 \right) U(Q) = E U(Q) \quad (\text{B.2.c})$$

El método que seguiremos se basa en construir los vectores propios de H por aplicación de operadores apropiados sobre uno de ellos. Puede ser mirado como un método de construcción del espacio vectorial E de los estados dinámicos del sistema.

B.3 INTRODUCCION DE LOS OPERADORES a, a^\dagger , y N .

Pongamos

$$a = \frac{\sqrt{2}}{2} (Q + i P) \quad (B.3.a)$$

$$a^\dagger = \frac{\sqrt{2}}{2} (Q - i P)$$

donde vemos que a y a^\dagger son hermíticos conjugados uno del otro. Verifiquemos: $(a, a^\dagger) = 1$ (B.3.b)

según puede comprobarse con solo sustituir y tener en cuenta la relación (B.2.b). Reemplazando Q y P por sus expresiones obtenidas de (B.3.a), en la ecuación (B.2.a)

$$H = \frac{1}{2} (a a^\dagger + a^\dagger a) \quad (B.3.c)$$

y poniendo $N = a^\dagger a$, gracias a (B.3.b) queda

$$H = N + \frac{1}{2} \quad (B.3.d)$$

verificándose las importantes relaciones (gracias a B.3.b)

$$Na = a (N-1)$$

$$Na^\dagger = a^\dagger (N+1) \quad (B.3.e)$$

El problema de los valores propios que tratamos de resolver es equivalente al de construir los vectores propios del operador N , función de a y a^\dagger .

Demostremos el importante Teorema:

TEOREMA: Si $|v\rangle$ es vector propio de N y \underline{v} el valor propio correspondiente

(i) necesariamente $v \geq 0$

(ii) si $v = 0$, $a|v\rangle = 0$; si no, $a|v\rangle$ es un vector no nulo de

norma

$$v < v | v \rangle$$

y es un vector propio de N correspondiente al valor propio $v - 1$.

(iii) $a^\dagger |v\rangle$ no es ciertamente nulo, su norma es

$$(v + 1) \langle v | v \rangle$$

y es un vector propio de N correspondiente al valor propio $v + 1$

DEMOSTRACION:

Por hipótesis $N |v\rangle = v |v\rangle$, $\langle v | v \rangle > 0$

Las normas de $a |v\rangle$ y de $a^\dagger |v\rangle$ son respectivamente

$$\langle v | a^\dagger a |v\rangle = \langle v | N |v\rangle = \langle v | v |v\rangle = v \langle v | v \rangle \quad (\text{B.3.f})$$

$$\langle v | a a^\dagger |v\rangle = \langle v | (N + 1) |v\rangle = (v + 1) \langle v | v \rangle \quad (\text{B.3.g})$$

La norma de todo vector del espacio de Hilbert es positiva o nula y su nulidad es una condición necesaria y suficiente para la nulidad del vector. Para que esto se verifique aquí será necesario y suficiente que $v \geq 0$ (propiedad (i)). La condición de nulidad de $a |v\rangle$ es un caso particular de la ecuación (B.3.F). Por otra parte $a |v\rangle$ y $a^\dagger |v\rangle$ verifican evidentemente las ecuaciones de valores propios anunciados ya que, según (B.3.e)

$$\begin{aligned} N a |v\rangle &= a (N - 1) |v\rangle = (v - 1) a |v\rangle \\ N a^\dagger |v\rangle &= a^\dagger (N + 1) |v\rangle = (v + 1) a^\dagger |v\rangle \end{aligned} \quad \text{c.q.d.}$$

B.4 ESPECTRO Y SISTEMA DE BASE DE N .

Si $v > 0$, el teorema precedente se aplica lo mismo al vector $a |v\rangle$ cuyo valor propio es $v - 1$. Esto nos asegura que $v \geq 1$. Si $v > 1$, el teorema se podrá volver a aplicar al vector $a^2 |v\rangle$. Se forma así, sucesivamente, la sucesión de vectores propios

$$a |v\rangle, a^2 |v\rangle, \dots, a^p |v\rangle, \dots$$

correspondiendo a los valores propios

$$v - 1, v - 2, \dots, v - p, \dots$$

Esta sucesión es de hecho limitada ya que los valores propios de N quedan limitados inferiormente por 0. Dicho de otra forma, los vectores de esta sucesión son todos nulos a partir de un cierto $n + 1$: la acción de a sobre el vector propio no nulo $a^n |v\rangle$, correspondiente al valor propio $v - n$, da 0; según (ii) esto exige que $v = n$.

De igual manera podemos aplicar el teorema al vector $a^{\dagger} |v\rangle$ que corresponde al valor propio $v+1$, después al vector $a^{\dagger 2} |v\rangle$ y así sucesivamente. Se forma así una sucesión ilimitada de vectores no nulos

$$a^{\dagger} |v\rangle, a^{\dagger 2} |v\rangle, \dots, a^{\dagger p} |v\rangle, \dots$$

vectores propios de N correspondientes a los valores propios

$$v+1, v+2, \dots, v+p, \dots$$

En conclusión, el espectro de valores propios de N está formado por la sucesión de números enteros no negativos. Además, por acción repetida de a o a^{\dagger} sobre uno de ellos, se obtiene una sucesión de vectores propios correspondiendo cada uno a uno de los valores propios del espectro. Este conjunto de vectores forma un sistema completo. En efecto, se puede mostrar que toda función de a y a^{\dagger} que conmuta con N es una función de N . Como consecuencia, N forma por si mismo un conjunto completo de observables que conmutan y ninguno de sus valores propios es degenerado.

Los vectores así contruidos no están normalizados a la unidad. Para formar una base ortonormada del observable N , es suficiente con multiplicar cada uno de ellos por una constante apropiada que se deduce de acuerdo con las relaciones (B.3.f-g) Se forma así la sucesión de vectores ortonormados

$$0, 1, \dots, n, \dots \quad (\text{B.4.a})$$

correspondiendo respectivamente a los valores propios siguientes de N

$$0, 1, \dots, n, \dots$$

deduciendose los unos de los otros por las relaciones de recurrencia

$$a^{\dagger} |n\rangle = (n+1)^{\frac{1}{2}} |n+1\rangle \quad (\text{B.4.b})$$

$$a |n\rangle = n^{\frac{1}{2}} |n-1\rangle \quad n \neq 0 \quad (\text{B.4.c})$$

$$a |0\rangle = 0 \quad (\text{B.4.d})$$

vemos que todos se deducen del vector $|0\rangle$ por la relación

$$|n\rangle = (n!)^{-\frac{1}{2}} a^{\dagger n} |0\rangle \quad (\text{B.4.e})$$

que obedecen efectivamente a la ecuación de valores propios

$$N |n\rangle = n |n\rangle \quad (\text{B.4.f})$$

y que su norma es igual a 1, verificando las relaciones

$$\langle n | n' \rangle = \delta_{nn'}$$

Ya que N forma el solo un conjunto completo, la sucesión de vectores (B.4.a) forma un sistema completo de vectores del espacio E de estados dinámicos del sistema cuántico considerado.

Solo nos resta ahora deshacer los cambios que fuimos haciendo

Como era $H = N + \frac{1}{2}$, los valores propios de H serán, evidentemente

$$\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots, n + \frac{1}{2}, \dots$$

y como hicimos $H = \hbar \omega H$, nos queda finalmente que los valores propios del Hamiltoniano del sistema, esto es, los niveles de energía del oscilador, serán

$$E_n = (n + \frac{1}{2}) \hbar \omega \quad n = 0, 1, 2, \dots$$

de donde vemos que quedan separados por intervalos de valor $\hbar \omega$ uno de otro. La energía del estado fundamental es $\frac{1}{2} \hbar \omega$, haciendo hincapié en que no es cero.

C. MOMENTO CINÉTICO O ANGULAR

C.1 DEFINICION DEL MOMENTO CINÉTICO

Por definición, el momento cinético \underline{l} de la partícula es

$$\underline{l} \equiv \underline{r} \times \underline{p}$$

siendo \underline{r} y \underline{p} respectivamente el vector de posición y de impulso de una determinada partícula. En mecánica cuántica, \underline{l} viene representado por el operador vectorial $-i\hbar (\underline{r} \times \nabla)$. Las tres componentes de este vector verifican las siguientes relaciones, que se obtendrán mediante las reglas de operación con conmutadores dadas anteriormente.

$$\begin{aligned} (l_x, l_y) &= (yP_z - zP_x, xP_z) \\ &= (yP_z, xP_z) + (zP_x, xP_z) \\ &= y(P_z, x)P_z + P_x(z, xP_z) \\ &= i\hbar (x \cdot P_y - y \cdot P_x) \\ &= i\hbar l_x \end{aligned}$$

2.1 GENERALIDADES, SEPARACION DE LOS MOVIMIENTOS ELECTRONICOS Y NUCLEARES.

Una molécula es un estado ligado constituido por la reunión de muchos átomos. Consiste que en un cierto número de núcleos atómicos - alrededor de los cuales evolucionan los electrones.

La búsqueda de los estados estacionarios de un sistema tan complejo es un problema extremadamente complicado. Nos presenta una característica que lo simplifica, la masa de los electrones es mucho más pequeña que la de los núcleos. Y en cambio, las fuerzas que actúan sobre ellos son aproximadamente de intensidad comparable. Como consecuencia, el movimiento de los electrones es mucho más rápido que el de los núcleos. Con esto, los dos movimientos pueden en una correcta aproximación ser tratados separadamente. En efecto, ^{ta} la primera aproximación, los electrones "ven" los núcleos como centros de fuerzas inmóviles y su estado dinámico el de un sistema de electrones en presencia de núcleos fijos; como estos se desplazan lentamente, el estado dinámico de los electrones sigue adiabáticamente esta evolución del potencial en el que se encuentran sumidos. En contrapartida, ya que los electrones ejecutan un gran número de revoluciones en cada desplazamiento apreciable de los núcleos, sobre estos no subsiste prácticamente mas que el efecto medio; para determinar el movimiento de los núcleos, se puede pues un gran error reemplazar la ^{ta} con los electrones por su valor medio en muchas resoluciones electrónicas, lo que da una ecuación de Schrödinger en la cual las variables de los electrones han desaparecido por completo. La aproximación sobre la cual se funda este método de separación de variables, lleva el nombre de aproximación adiabática.

Por otra parte, el potencial que figura en la ecuación de Sch. de los núcleos, depende únicamente de las distancias mutuas entre ellos. Ya que la molécula existe, este potencial presenta necesariamente un mínimo para unos valores bien determinados y finitos de las distancias internucleares; este mínimo corresponde al punto de equilibrio estable del sistema, punto alrededor del cual los núcleos podrán efectuar sus pequeñas oscilaciones. A estas vibraciones internas de los núcleos pueden superponerse movimientos de traslación y de rotación del conjunto. El movimiento de traslación se separa por completo y de una manera rigurosa del resto por la introducción del centro de masas; se sabe que esto último se desplaza como una partícula libre con masa igual a la total del sistema. En la discusión que sigue, supondremos hecha esta separación del centro de masa y no nos fijemos mas que en los movimientos de rotación y de vibración de los núcleos.

2.2 MOVIMIENTO DE LOS ELECTRONES EN PRESENCIA DE NUCLEOS FIJOS

Vamos a tratar de encontrar ahora los estados estacionarios de la molécula, efectuando unas simplificaciones que abrevien el cálculo pero sin que afecten al principio mismo del método. En particular trataremos a los electrones y a los núcleos como partículas sin spin.

Convendremos en representar los grados de libertad de electrones y núcleos por los índices i y j respectivamente; designaremos por x_i y X_j las coordenadas de los electrones y los núcleos de masas m_i y M_j respectivamente. Sea T_e la energía cinética de los electrones T_n la de los núcleos y V el potencial de interacción entre las diversas partículas que componen la molécula. El Hamiltoniano H de la molécula es así la suma de estos tres términos.

$$H = T_e + T_n + V \quad (2.2.a)$$

donde

$$T_n = - \sum_j \frac{\hbar^2}{2M_j} \frac{\partial^2}{\partial X_j^2} ; T_e = - \sum_i \frac{\hbar^2}{2m_i} \frac{\partial^2}{\partial x_i^2} \quad (2.2b.)$$

$V \equiv V(x, X)$ es una cierta función de las coordenadas de los electrones y de los núcleos: es la suma de las energías coulombianas de las partículas del sistema tomadas dos a dos.

Consideremos el sistema mas simple:

$$H^{(0)} = T_e + V$$

que se obtiene despreciando la energía cinética de los núcleos en la expresión de H . $H^{(0)}$ es el Hamiltoniano del sistema en el límite de $M_j \rightarrow \infty$; sus estados estacionarios son los del sistema de electrones en presencia de los núcleos fijos.

Con esto como se verifica que $H^{(0)}$ no contiene derivadas con respecto a las X_j y se cumple para toda función $A(p, q)$ que

$$(q, A(p, q)) = i \hbar \frac{\partial A}{\partial p}$$

se tendrá evidentemente $(x_j, H^{(0)}) = 0$ (2.2b.)

lo que implica la presencia mutua de ambas cantidades. De ello resulta que para resolver el problema de los valores propios de $H^{(0)}$ puede fijarse unos valores determinados de X'_j para las coordenadas de posición de los núcleos y buscar los vectores propios de $H^{(0)}$ entre los vectores correspondientes a estos valores precisos de X_j . Designemos por X' el conjunto de los X'_j . A cada conjunto X' dado corresponde un conjunto $W_n(X')$ de valores propios de $H^{(0)}$; el número cuántico n servirá para distinguir a unos de otros valores de este conjunto. Para n y X' dados, se tendrán uno o eventualmente varios vectores propios linealmente independientes; en este último caso, se los podía distinguir a unos de otros por un nuevo índice s . Con esto, la ecuación de Schrödinger de $H^{(0)}$ será:

$$H^{(0)} |n s X'\rangle = W_n(X') |n s X'\rangle \quad (2.2.c)$$

En la representación (x, X) , el vector propio $|n s X'\rangle$ puede ser representado por la función de onda

$$\psi_{ns}(x, X') \quad (X=X')$$

y será la solución de la ecuación de Schrödinger

$$(T_e + V(x, X')) \psi_{ns}(x, X') = W_n(X') \psi_{ns}(x, X') \quad (2.2.e)$$

En la ecuación (2.2.e), X' juega el papel de un simple parámetro es la ecuación de Schrödinger de los electrones de la molécula cuando se fija en X' las posiciones de sus núcleos. A cada solución propia esta ecuación corresponde una función propia $H^{(0)}$ de la forma (2.2.d); conociendo todas las soluciones propias de la ecuación (2.2.e) para todos los valores posibles de X' , se forma así un sistema completo de funciones de $H^{(0)}$.

El problema de los valores propios de (2.2.e) es análogo al de la búsqueda de los estados de un átomo; puede ser resuelto por métodos análogos. En el caso presente, en cambio, los electrones se encuentran sometidos a varios centros de fuerzas y la simetría encontrada en los problemas atómicos se encuentra aquí totalmente destruida.

Veamos algo de esto de la simetría. El potencial $V(x, X)$ es invariante por traslación, rotación y reflexión del conjunto de electrones y núcleos; es además invariante por permutación de las partículas del sistema.

Tomas las simetrías que pueda presentar $H^{(0)}$ son evidentemente las mismas que las de $V(x, X)$. La degeneración del nivel $W_n(X')$ vendrá ligado a estas propiedades de simetría.

Convendrá pues estudiar de una forma mas amplia, las posibles simetrías de una molécula y su repercusión sobre la degeneración de los correspondientes niveles energéticos y las estados con ella ligados. Esto será lo que veremos en el tema III.

III - TEORIA DE LA SIMETRIA

Transformaciones simétricas

- Simetría de un cuerpo: cto de cambios que le dejan invariable.
- Toda simetría puede descomponerse en suma de tres: rotación en torno a un eje; reflexión en relación a un plano; desplazamiento paralelo. El último solo a medio ∞ : estructura cristal. Una molécula solo: rotación y reflexión.
- Eje de simetría de orden n : el cuerpo queda inalterado por un giro de $2\pi/n$ en torno a él. Se designa por C_n . Rotaciones hechas en i veces es C_n^i . Si $n/i = n^\circ$ entero $C_n^i = C_{n/i}$. En particular $C_n^n = E$.
- Plano de simetría: El cuerpo queda inalterado por reflexión según un plano. Se designa por σ_h . Se verifica $\sigma_h^2 = E$.
- Eje de rotación-reflexión: aplicación simultánea de rotación según un eje y reflexión según un plano perpendicular a él. Un sólido tiene un eje de este tipo de orden n cuando un giro de $2\pi/n$ seguido de una reflexión, la deja invariable. Solo en forma de simetría - cuando $n = \text{par}$. Si es impar todo se reduce a un eje de orden n y un plano de simetría. Se designa por S_n . Si es σ_h la reflexión, será $S_n = C_n \sigma_h = \sigma_h C_n$

Caso particular: $S_2 = C_2 \sigma_h \equiv I$ inversión. El cuerpo tiene centro de simetría.

Se verifica: $I \sigma_h = C_2 = \sigma_h \Rightarrow$ la existencia de dos obliga la del tercero.

- Propiedades geométricas: El producto de dos rotaciones según ejes que se cortan es otra rotación según un tercer eje que pasa por el punto de corte.
El producto de dos reflexiones según planos que se cortan equivale a una rotación con eje la línea de corte de los dos planos y ángulo el doble del que forman los planos. Si es $C(\theta)$ la rotación según C con ángulo θ y las reflexiones σ_v y σ_v' es $\sigma_v \sigma_v' = C(2\theta)$ con θ como ángulo entre los planos. Interviene el orden de la operación: $(\sigma_v, \sigma_v') \neq 0$; la rotación resultante es de sentido contrario. Es $\sigma_v' = \sigma_v C(2\theta)$, o sea, una rotación seguida de una reflexión equivale a otra reflexión. De aquí que si hay un eje C_2 y dos planos de simetría que pasan por el y son \perp entre si, son mutuamente dependientes: dos \Rightarrow el tercero.

Grupos de transformaciones.

Las transformaciones simétricas son transformaciones de las coordenadas que dejan invariante al Hamiltoniano del sistema. Al no cambiar el sistema por rotación o reflexión, por ellas permanece invariable la ecuación de Schrödinger. Los grupos de simetría se estudian por teoría de grupos. Elementos de la teoría de grupos: propiedades de grupo. Grupo Abelian. Grupo cíclico: orden de un elemento; periodo de un elemento (A). Orden de un grupo. Elementos conjugados, $A = C B C^{-1}$. Clases de grupo o grupo cociente: subcto de elementos conjugados; no es subgrupo (no posee el E).

Grupos puntuales

Aquellos que todas sus transformaciones dejan fijo un punto del cuerpo.

Solo enumeramos aquellos de interés para el estudio de las moléculas mas comunes.

- I- Grupos C_n : eje de simetría de orden n . Cíclico. Cada uno de los elementos forma una clase por si mismo. $C_1 = E$.
- II- Grupos S_{2n} : rotaciones-reflexiones en torno a eje de orden $2n$. $2n$ elementos. Cíclico. $S_2 = (E, I) = C_1$. Si $2n = 4p + 2$ tiene la inversión entre sus elementos ya que $(S_{4p+2})^{2p+1} = C_2 \sigma_h = I$

III- Grupos C_{nh} : se obtiene adicionando un plano de simetría perpendicular a un eje C_n . Tiene $2n$ el n rotaciones y n rotaciones-reflexiones $C_n^k \sigma_h$, $k = 1, 2, \dots, n$ (incluye la reflexión $C_n^n \sigma_h = \sigma_h$).

El grupo es abeliano. El número de clases es igual al de elementos. Si $n = 2$ p posee centro de simetría ($C_{2p}^p \sigma_h = C_2 \sigma_h = I$). El mas simple es C_{1h} contiene E y σ_h ; se designa por C_s .

IV- Grupos C_{nv} : por adición a un eje C_n de un plano de simetría que pasa por el. Aparecen otros $n-1$ planos, que pasan también por el eje y forman ángulos π/n . Es por ser $\sigma_v = C(2\pi/n) \sigma_v$. Contiene $2n$ elementos: n rotaciones y n reflexiones según los σ_v . Si n es impar hay $p + 2$ clases, si n es par, hay $p + 3$.

V- Grupos D_n : por adición a un eje de simetría de orden n de un eje de segundo orden perpendicular a el; aparecen con esto, $n-1$ ejes posteriores de tal tipo, por lo que hay en total n ejes horizontales de segundo orden y formando ángulos de π/n . El grupo resultante contiene $2n$ elementos: n rotaciones en torno al eje de orden n y n rotaciones en torno a los ejes horizontales de valor π (designaremos a estos últimos por σ_v . Solo es C_2 cuando sea eje vertical). Contiene $p + 3$ clases si es n par y $p + 2$ si es impar.

VI- Grupo D_{nh} : por adición al anterior de un plano horizontal de simetría que pase por los n ejes horizontales, aparecen n planos verticales que pasan por el eje vertical y cada uno de los horizontales. Contiene $4n$ elementos; aparte de los $2n$ elementos del D_n , contiene n reflexiones σ_v y n rotaciones. El número de clases será también doble que el de D_n .

VII- Grupos D_{nd} : por adición de planos verticales de simetría a los ejes del grupo D_n , en el centro de los ejes binarios de simetría. La adición de un plano de este tipo supone la posterior aparición de otros $n-1$ planos. Contiene $4n$ elementos; los $2n$ del grupo D_n , n reflexiones reflexiona en los planos verticales y n transformaciones de la forma $G = U_2 \sigma_d$, que no son otra cosa que $U_2 = \sigma_h \sigma_v$ donde σ_v es una reflexión en el plano vertical que pasa a través del correspondiente eje de segundo orden, entonces $G = \sigma_h \sigma_v \sigma_d = \sigma_h C_{2n}^{2k+1} = S_{2n}^{2k+1}$ esto es un eje de rotacion-reflexión de segundo orden. Contiene $2p + 3$ clases.

Aparte de los grupos vistos, existen otros que aquí no consideramos ya que las moléculas que normalmente se encuentran no poseen una simetría tan complicada. Son los grupos de tetraedro T , el T_d , el T_h , el del octaedro O , el O_h y los del icosaedro Y , Y_h .

3.4 REPRESENTACION DE GRUPOS.

3.4.1 DEFINICIONES

Grupos de sustituciones lineales

El producto de matrices cuadradas es una operación asociativa. Si un conjunto de matrices $n \times n$ satisface los axiomas de definición de un grupo. Formará un cierto grupo (G).

Cada matriz representa un cierto operador lineal G de un espacio vectorial E_n de n dimensiones y define así una transformación lineal de los vectores de este espacio: si $1, 2, \dots, n$ son los n vectores de base de este espacio (en principio esta base no tendrá porque ser ortonormada), la transformación de cada uno de ellos vendrá dado por la ecuación:

$$G \quad K \quad = \quad \sum_j G_{jk} \quad j$$

El tipo de grupo que acaba de ser definido lleva el nombre de grupo de sustituciones lineales, de n dimensiones.

Representación de un grupo.

Por definición, una representación (lineal) de un grupo G es un grupo de sustituciones lineales con las cuales G es homomorfo.

Sea (G) un grupo de este tipo y E el espacio vectorial sobre el cual actúan sus matrices. Por definición, E es el espacio de representación y el número n de sus dimensiones, el grado de la representación.

Si G es isomorfo a (G) , la representación es fiel.

Representaciones de grado 1.

Todo grupo posee al menos una representación de grado 1, su representación idéntica, en la cual cada elemento del grupo viene representado por el número 1.

Representación unitaria.

La representación (G) es unitaria si todos sus matrices son unitarios.

Representaciones equivalentes.

Dos representaciones (G) y (G') son equivalentes si tienen el mismo número de dimensiones y si cada matriz $G'(g)$ de la una se deduce por una cierta transformación lineal T de la matriz $G(g)$ de la otra representado la misma operación del grupo g :

$$G'(g) = T G(g) T^{-1} \quad \forall g \in G$$

y en general $(G') = T (G) T^{-1}$

pudiendo escribirse simbólicamente $(G') \simeq (G)$

Si se identifican los espacios de representación E y E' , el paso de G a la representación equivalente (G') corresponde a un cambio de los vectores de la base en este espacio.

Representaciones conjugadas.

Por definición, las dos representaciones $(G), (G^*)$, en las que las matrices $G(g), G^*(g)$ son complejas conjugadas una de la otra, son conjugadas.

Una representación (G) es autoconjugada si su conjugada le es equivalente

$$(G) \simeq (G^*)$$

Caracteres.

Por definición, la traza de la matriz $G(g)$ representante de la operación g es la representación (G) de g es el caracter X de g en esta representación

$$X(g) \equiv \text{Tr } G(g)$$

Según las propiedades de la traza, dos operaciones pertenecientes a la misma clase, tienen el mismo caracter: el caracter es una función de clase.

Demostración

$$X'(g \cdot) = \sum_i G'_{ii} = \sum_{i,k,e} S^{-1}_{ik} G_{ke} S_{ei} = \sum_{ke} G_{ke} \delta_{ke} = \sum_k G_{kk} = X(g)$$

Por la misma razón, dos representaciones equivalentes tienen el mismo sistema de caracteres

$$X'(g) = X(g) \quad \text{si} \quad G'(g) = T G(g) T^{-1}$$

se escribe simbólicamente

$$X' = X \quad \text{si} \quad (G') \sim (G)$$

3.4.2 OPERACIONES SOBRE LOS ESPACIOS DE REPRESENTACION. REDUCTIBILIDAD. SUMA DIRECTA.

Sean (G^a) , (G^b) dos representaciones de un mismo grupo G , n_a, n_b sus grados respectivos, E_a, E_b sus respectivos espacios de representación.

Si $|a_1\rangle, |a_2\rangle, \dots, |a_n\rangle$ son los vectores de base de E_a , $|b_1\rangle, |b_2\rangle, \dots, |b_n\rangle$ los de E_b las sustituciones lineales que representan la operación g en estas dos representaciones vendrán definidas respectivamente por las leyes de transformación de los vectores de base:

$$g(|a_k\rangle) = \sum_{\lambda} |a_{\lambda}\rangle G_{\lambda k}^a(g) \quad g(|b_{\mu}\rangle) = \sum_{\nu} |b_{\nu}\rangle G_{\nu \mu}^b(g) \quad (3.4.2.a)$$

Por definición, la suma directa $E_a \oplus E_b$ de los espacios E_a y E_b es el espacio subtendido por la $n_a + n_b$ vectores $|a_1\rangle, \dots, |a_{n_a}\rangle, |b_1\rangle, \dots, |b_{n_b}\rangle$.

Las matrices de este nuevo espacio pueden ponerse bajo la forma

$$M = \begin{pmatrix} M_{aa} & M_{ab} \\ M_{ba} & M_{bb} \end{pmatrix}$$

M_{aa} es una matriz $n_a \times n_a$ transformando los vectores, del espacio E_a unos en otros; M_{bb} es una matriz $n_b \times n_b$ transformando los del E_b ; M_{ab} es una matriz $n_a \times n_b$ que transforma los vectores del espacio E_b en vectores del espacio E_a ; M_{ba} es una matriz $n_b \times n_a$ que posee la propiedad inversa. En particular si A es una matriz de E_a , B otra de E_b , se puede construir por suma directa la matriz $A \oplus B$ que posee una forma de bloques.

$$A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

se ve inmediatamente que

$$T_r(A \oplus B) = T_r A \oplus T_r B$$

$$\det(A \oplus B) = \det A \det B \quad (3.4.2.b)$$

La operación de suma directa sobre las matrices conserva la unidad y la hay de multiplicación de matrices

$$1_{(a+b)} = 1_{(a)} \oplus 1_{(b)} \quad (A_1 \oplus B_1)(A_2 \oplus B_2) = A_1 A_2 \oplus B_1 B_2$$

Como consecuencia el conjunto (G^{a+b}) de las matrices $G^a(g) \oplus G^b(g)$ constituye una representación del grupo G . La sustitución lineal que representa la operación g está definida por las leyes (3.4.2.a) de transformaciones de los $n_a + n_b$ vectores de base de $E_a \oplus E_b$. En cuanto a los caracteres de esta representación, se tiene:

$$X^{a+b}(g) = X^a(g) \oplus X^b(g), \text{ o dicho de otra forma } X^{a+b} = X^a \oplus X^b$$

REDUCTIBILIDAD.— Un subespacio invariante del espacio E de una representación (G) es un subespacio de E en el que los vectores del mismo se transforman los unos en los otros por la aplicación de las matrices de (G) .

La representación G se denomina:

(i) irreducible si E no tiene ningún subespacio invariante (aparte de el mismo y del espacio nulo).

(i i) reducible en el caso contrario

$$E = E_1 \oplus E_2 \quad (E_1, E_2 \neq 0), E, \text{ subespacio invariante.}$$

En el segundo caso, si E_2 es igualmente invariante, (G) se dice descomponible. Se puede entonces transformar por una transformación lineal apropiada, los vectores de Base E en vectores bien situados en E_1 o en E_2 . La representación así obtenida es la suma directa de una representación (G_1) de espacio E_1 y de otra representación (G_2) de espacio E_2 :

$$(G) \simeq (G_1) \oplus (G_2)$$

(G_1) y (G_2) llevan el nombre de componentes de (G)

Una representación (G) es componente reductible si se la puede poner bajo la forma de una suma directa de componentes reductibles:

$$(G) \simeq (G^{(1)}) \oplus (G^{(2)}) \oplus \dots \oplus ((G^{(1)}), (G^{(2)}) \dots \text{irreducible})$$

Toda representación unitaria es bien irreducible o bien completamente reductible.

Todas las representaciones irreducibles de un grupo abelino son de grado 1.

APLICACIÓN HOMOMORFICA DE UN ESPACIO DE REPRESENTACION SOBRE OTRO.— Se ha realizado una aplicación lineal de E_a en E_b si se hace corresponder a cada vector $|a\rangle$ de E_a uno y uno solo vector $|b\rangle$ de E_b de tal forma que esta correspondencia sea lineal. Esta aplicación es homomorfica si conserva las propiedades de transformación de los vectores en las diversas operaciones del grupo, es decir si la correspondencia

$$|a\rangle \rightarrow |b\rangle \quad \text{entraña la correspondencia} \quad g(|a\rangle) \rightarrow g(|b\rangle) \quad \forall g \in G$$

Una aplicación de E_a sobre E_b está completamente determinada si se conoce la matriz $n_b \times n_a$ que define el vector de E_b correspondiente a cada vector de base de E_a :

$$|a_x\rangle \rightarrow \sum_n |b_n\rangle S_{ux} \quad (3.4.2.c)$$

Si la aplicación es homomorfica, se tiene la ecuación entre matrices

$$S G^a(g) = G^b(g) S \quad \forall g \in G$$

$$\text{esto es} \quad S (G^a) = G^b S \quad (3.4.2.d)$$

Cuando el conjunto de los vectores $|b\rangle$ correspondiente a los vectores de E_a subtiende la totalidad de E_b (esto supone que $n_a \geq n_b$) se tiene una aplicación de E_a sobre E_b . En este caso, todos los determinantes $n_b \times n_b$ contenidos en S son diferentes de cero.

Cuando además esta correspondencia es binnivoca ($n_a = n_b$), la matriz S es no singular: $\det S \neq 0$. En este caso, la aplicación homomorfica de E_a sobre E_b lleva el nombre de correspondencia isomorfica de y se tiene evidentemente $(G^a) \simeq (G^b)$

PRODUCTO TENSORIAL O KRONECKERIANO.-

Producto tensorial de las matrices.

El producto tensorial de dos matrices de ordenes N_1 y N_2 respectivamente es una matriz de orden $N_1 N_2$. En particular, el producto tensorial de las matrices unidad $I^{(1)}$ y $I^{(2)}$ es una matriz unidad $I^{(12)}$ de $N_1 N_2$ dimensiones.

Se puede mirar estas matrices productos tensoriales como las matrices de uno de los espacios en las que cada elemento de las mismas es una matriz del otro multiplicada por el número que ocupará ese lugar.

Ej: Sea

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

$$A \otimes B = \begin{pmatrix} a_{11} B & a_{12} B \\ a_{21} B & a_{22} B \end{pmatrix} \text{ que es pues de dimensión } 2 \times 3 = 6$$

Se conviene en representar las líneas (y las columnas por dos índices m, n de los que el primero se refiere a las componentes de uno de los espacios y el segundo a las del otro, las líneas (y las columnas) se encuentran en el orden: 11, 12, 21, 22. (si se trata de productos 2×2).

Por combinación lineal de las matrices productos tensoriales, se forman matrices cuadradas de dos índices

$$A_{m_1 m_2; n_1 n_2} \quad (m_1, n_1 = 1, 2, \dots, N_1; m_2, n_2 = 1, 2, \dots, N_2)$$

y de $N_1 N_2$ dimensiones. Por definición, la traza parcial en uno de los espacios es

$$(T_{r_1} A)_{m_2 n_2} \equiv \sum_{n_1=1}^{N_1} A_{n_1 m_2; n_1 n_2}$$

Es evidente que

$$T_r A = T_{r_2} (T_{r_1} A) = T_{r_1} (T_{r_2} A)$$

y que si $A = A^{(1)} \otimes A^{(2)}$

$$T_r A = T_r (A^{(1)} \otimes A^{(2)}) = (T_{r_1} A^{(1)}) (T_{r_2} A^{(2)})$$

Producto tensorial de espacios.-

Efectuando el producto tensorial del espacio E_a por el E_b se forma el espacio $E_a \otimes E_b$ de $n_a n_b$ dimensiones en el que sus vectores base son $|a b x u\rangle \equiv |a x\rangle \otimes |b u\rangle$ ($x = 1, \dots, n_a$) ($u = 1, \dots, n_b$). Las matrices $G^a(g) \otimes G^b(g)$ formadas por el producto tensorial de las matrices que representan g en (G^a) y (G^b) forman una representación $(G^{ab}) = (G^a) \otimes (G^b)$ de grado $n_a n_b$ del grupo G , en el cual una operación dada de este grupo está definida por la ley

$$g(|a b x u\rangle) = \sum_{\lambda, \mu} |a b \lambda \mu\rangle G_{\lambda x}^a(g) G_{\mu u}^b(g)$$

Los caracteres de esta representación vienen dados por las relaciones

$$\chi^{ab}(g) = \chi^a(g) \chi^b(g)$$

Caso particular del anterior es cuando las dos representaciones (G^a) y (G^b) son la misma. Evidentemente se verifica aquí que

$$X^{aa}(g) = (X^a(g))^2$$

Siendo la representación $(G^{aa}) = (G^a) \times (G^a)$ de grado n_a^2 y formado el espacio de representación por n_a^2 vectores base de la forma

$$|a a u x\rangle \equiv |a u\rangle |a x\rangle \quad (x, u = 1, \dots, n_a)$$

Esta nueva representación puede ser descompuesta en dos representaciones de menores dimensiones. Una de ellas viene dada por las $\frac{1}{2}n(n+1)$ funciones $|a i\rangle |a k\rangle + |a k\rangle |a i\rangle$ y las $\frac{1}{2}n(n-1)$ funciones $|a i\rangle |a k\rangle - |a k\rangle |a i\rangle$ $i \neq k$. Es evidente que las funciones de uno de estos conjuntos se transforman solamente en combinaciones de ellas mismas.

LEMA DE SCHUR.-

Si (G^a) y (G^b) son dos representaciones irreducibles de un mismo grupo y si existe una aplicación homomórfica del espacio de la una en el espacio de la otra, la matriz S que define esta aplicación posee las propiedades siguientes

- a) si (G^a) y (G^b) son inequivalentes, necesariamente $S = 0$
- b) si $(G^a) \simeq (G^b)$, o bien $S=0$, o bien $\det S \neq 0$
- c) si $(G^a) \equiv (G^b)$, S es un múltiplo de la matriz unidad:

$$S = c \frac{1}{2} (C.cte.)$$

REPRESENTACIONES COMPLETAMENTE REDUCTIBLES.T.-

Sean dos descomposiciones en partes irreducibles de una representación completamente reductible (G) :

$$(G) \simeq (G_1) \oplus (G_2) \oplus \dots \oplus (G_p) \quad (G) \simeq (G'_1) \oplus (G'_2) \oplus \dots \oplus (G'_{p'})$$

se puede demostrar que $p = p'$ y que se puede hacer corresponder de manera biunívoca a cada término de la primera composición en término de la segunda que le es equivalente. Dicho de otra forma:

TEOREMA DE UNICIDAD.

Si (G) es completamente reductible, su descomposición en partes irreducibles es única.

A partir de aquí, salvo indicaciones en contra, no consideraremos como distintas dos representaciones equivalentes una a la otra.

Así, según el teorema de unicidad, toda representación completamente reductible (G) obedece a la ley de equivalencia

$$(G) \simeq \sum_j n_j (G^{(j)})$$

y la sucesión de enteros (≥ 0) $n_1, n_2, \dots, n_j, \dots$ definida de manera única. Igualmente, el sistema de caracteres de (G) obedece a la ecuación

$$X = \sum_j n_j X^{(j)}$$

REPRESENTACIONES IRREDUCTIBLES.-

a) Número: el número de representaciones irreducibles inequivalentes es igual al número L de clases

b) Grado: Si d_j es el grado de la j representación irreducible $F^{(j)}$ ($j = 1, 2, \dots, L$)

N/d_j es entero

$$\sum_{j=1}^L d_j^2 = N$$

c) Relaciones de ortogonalidad: Si las representaciones irreducibles unitarias $F^{(j)}$, $F^{(k)}$ son bien inequivalentes, o bien iguales

$$\frac{d_j}{N} \sum_{\alpha=1}^N (F | j \alpha \beta) (F | k \gamma \delta)^x = \delta_{jk} \delta_{\alpha\gamma} \delta_{\beta\delta}$$

de donde se obtiene la relación de ortogonalidad de los caracteres

$$\sum_{a=1}^L \frac{l_a}{N} (a | j) (a | k)^x = \delta_{jk}$$

TABLAS DE CARACTERES.-

Es conveniente representar los caracteres de las diversas representaciones en una "tabla de caracteres" para el grupo dado. Las columnas vienen encabezadas por las diversas clases, precedidas del número l_k de elementos en dicha clase. Las filas vienen precedidas por las representaciones irreducibles. Los elementos de la tabla son los $X^{(j)} G^{(k)}$. Estas tablas dan, según veremos, menor información acerca del grupo que la que nos daría el conjunto completo de matrices de las representaciones irreducibles; pero esta información es suficiente para la mayor parte de los propósitos que buscamos en ellas, y además reúne la peculiaridad, ya vista cuando dimos la teoría general de la representación que el caracter es una característica común a todas las representaciones equivalentes.

Como ejemplo de tabla de caracteres, y antes de entrar de lleno en su formación general, consideraremos el siguiente ejemplo.

Sean las siguientes matrices

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad B = \begin{pmatrix} -\frac{1}{2} & \sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix} \quad C = \begin{pmatrix} -1/2 & -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{3}}{2} & 1/2 \end{pmatrix} \quad D = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}$$

$$F = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}$$

que forman grupo y cuya tabla de multiplicación puede comprobarse es:

	E	A	B	C	D	F
E	E	A	B	C	D	F
A	A	E	D	F	B	C
B	B	F	E	D	C	A
C	C	D	F	E	A	B
D	D	C	A	B	F	E
F	F	B	C	A	E	D

que pueden tomarse como representación de las operaciones de simetría de un triángulo equilátero con A,B y C como rotaciones de π alrededor de los ejes mostrados, D como rotación de $2\pi/3$ en sentido de las agujas de un reloj y en torno a un eje que pase por un centro y F otra rotación análoga, pero de sentido contrario.

Pues bien, otra posible representación del mismo grupo puede ser obtenida por ejemplo tomando el determinante de cada matriz. Con esto se reducen las matrices a simples ± 1 , dando una representación unidimensional. Esta representación no es, evidentemente fiel ya que hay únicamente dos diferentes matrices mientras hay seis elementos en el grupo del que es representación.

Finalmente, la representación mas trivial, y a la que ya hemos aludido antes, será lo que a cada elemento del grupo le corresponde el mismo número: la unidad, representación idéntica, como se recordará.

Pues bien, vamos a representar la tabla de caracteres de estas tres representaciones que hemos encontrado. Según se puede comprobar el grupo del que estamos haciendo la representación, no es otro que el D_3 : un eje de simetría terciario y tres binarios. Sus elementos eran: E, C_3 , U_2 . Su número de clases: $E \rightarrow 1$; $C_3 \rightarrow 2$; $U_2 \rightarrow 3$. Tomando los caracteres de la representación primera, queda en total:

	E	$2C_2$	$3C_2$
G^1	1	1	1
G^2	1	-1	1
G^3	2	0	-1

Según puede comprobarse entre los elementos de las filas se verifica la relación expuesta anteriormente entre caracteres $\sum_{a=1}^L 1_a (a|j)(a|k)^* = \delta_{jk} N$. Y

también que las columnas forman vectores ortogonales. Esto no es acci-
dente. Puede demostrarse que es cierto en general.

CONSTITUCION DE LAS TABLAS DE CARACTERES.-

Aunque dadas de una manera general anteriormente, repetimos aquí las reglas de constitución de las tablas.

1. El número de representaciones irreducibles es igual al número de clases de elementos del grupo.

2. Las dimensiones d_j de las representaciones irreducibles quedan determinadas por el hecho de que $\sum_j d_j^2 = N$. En la mayor parte de los casos tiene una única solución. Ya que el elemento identidad debe ser representado por una matriz unidad, la traza de una matriz de la clase identidad es simplemente d_j . Esto determina la primera columna de la tabla, $X^{(j)}(E) = d_j$. También, como siempre tendremos la representación unidimensional en que cada grupo está representado por la unidad podremos escribir siempre la primera columna como $X^{(1)}(g^k) = 1$.

3. Las filas de la tabla deben ser ortogonales y normalizadas a N , con un factor de peso 1_x , o número de elementos en g^k . Esto es:

$$\sum_k X^{(i)}(g^k)^* X^{(j)}(g^k) 1_k = N \delta_{ij}$$

4. Las columnas de la tabla deben ser vectores ortogonales normalizados a $\frac{h}{1_k}$. Esto es

$$\sum_i X^i(g^k)^* X^i(g^e) = \frac{N}{1_k} \delta_{ke}$$

5. Los elementos dentro de la fila están relacionados por

$$1_i X^{(i)}(g^j) 1_k X^{(i)}(g^k) = d_i \sum_e c_{jke} 1_e X^i(g^e)$$

donde c_{jke} son constantes definidas por la expresión que gobierna la multiplicación de clases $g^j g^k = \sum_e c_{jke} g^e$

Las primeras tres reglas serán usualmente suficientes para obtener la tabla de caracteres, pero las dos últimas facilitan la inspección.

En el caso de que los caracteres no fueran números enteros, el procedimiento es menos simple.

DESCOMPOSICION DE REPRESENTACIONES REDUCIBLES.--

Como ya hemos visto, el caracter de una representación reductible G es la suma de los caracteres de las representaciones componentes irreducibles $G^{(j)}$. Podemos escribir

$$X(R) = \sum_j a_j X^{(j)}(R)$$

donde X es el caracter de R y a_j es el número de veces que $G^{(j)}$ aparece en G . Ya que $X^{(j)}(R)$ forma un sistema de vectores ortogonales, los coeficientes del desarrollo a_j pueden ser obtenidos como es usual mediante el producto escalar con $X^{(j)}(R)$. Así queda

$$\sum_R X(R) X^{(i)}(R)^* = \sum_R \sum_j a_j X^{(j)}(R) X^{(i)}(R)^* = N a_j$$

y de aquí

$$a_j = N^{-1} \sum_R X^{(j)}(R)^* X(R) = N^{-1} \sum_k l_k X^{(j)}(G^k)^* X(G^k)$$

REPRESENTACION REGULAR.⊙

Dada la tabla de multiplicación de un grupo, podemos formar siempre una representación reducible denominada representación regular como sigue: Escribir la tabla de multiplicación y reagruparla de nuevo las filas de forma que corresponda a los inversos de los correspondientes en las columnas. De esta manera se obtiene solamente el elemento E en la diagonal principal. La matriz de la representación regular para el elemento R del grupo se obtiene reemplazando R por la unidad y los demás por cero en la tabla resultante. Como ejemplo reagruparemos la tabla vista antes:

	E	A	B	C	D	F
E^{-1}	E	A	B	C	D	F
A^{-1}	A	E	D	F	B	C
B^{-1}	B	F	E	D	C	A
C^{-1}	C	D	F	E	A	B
D^{-1}	F	B	C	A	E	D
F^{-1}	D	C	A	B	F	E

$$G^{\text{res}}(A) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

Evidentemente, en general $X^{\text{reg}}(E) = N$ y $X^{\text{reg}}(R) = 0 \quad \forall R \neq E$ ya que por construcción G^{reg} solo tiene elementos no cero en la diagonal la correspondiente a E.

TEOREMA: La representación regular contiene cada representación irreducible un número de veces igual a la dimensión de la representación irreducible.

REPRESENTACIONES IRREDUCIBLES DE GRUPOS PUNTUALES

GRUPOS PUNTUALES.

Un grupo puntual es un grupo de rotaciones en torno a un punto, esto es, de rotaciones en torno a ejes que intersectan en un punto. Esto los distingue de los grupos espaciales más generales que relacionan traslaciones y rotaciones según ejes que no intersectan. Daremos aquí unas consideraciones sobre sus aspectos mas característicos, su manejo y concluiremos con algunos de ellos tabulados en las correspondientes tablas de caracteres.

REPRESENTACIONES IRREDUCIBLES.

Dado el número de elementos de cada grupo puntual y su estructura de clases, el número y las dimensiones de sus representaciones irreducibles quedan fijadas por las reglas dadas anteriormente. Con ellas pueden construirse así mismo las correspondientes tablas de caracteres. Esto ha sido hecho por numerosos autores y se adjunta una serie de ellas. Hay, según se ha demostrado, 32 grupos que no han de ser tabulados todos ya que muchos son producto directo de grupos mas simples con el grupo de la inversión.

Así, por ejemplo, el grupo D_{3h} está formado por el producto directo del D_3 y el de inversión $S = (E, n)$, esto es $D_{3h} = D_3 \times S$. Las tablas de estos dos últimos son:

S	E	n	D ₃	E	2C ₃	3C' ₂
A _g	1	1	A ₁	1	1	1
A _u	1	-1	A ₂	1	1	-1
			E	2	-1	0

multiplicándoles en la forma vista anteriormente para el producto tensorial queda.

D_{3h}	E	$2C_3$	$3C'_2$	σ_h	$2S_3$	$3\sigma_v$	Grupos isomorfos, como es lógico, tendrán la misma tabla de caracteres.
A'_1	1	1	1	1	1	1	
A'_2	1	1	-1	1	1	-1	
E'	2	-1	0	2	-1	0	
A''_1	1	1	1	-1	-1	-1	En la tabla anterior
A''_2	1	1	-1	-1	-1	1	pueden verse las diferentes representaciones
E''	2	-1	0	-2	1	0	irreducibles representadas

con unas ciertas letras. El convenio que se sigue es el siguiente: Las representaciones irreducibles monodimensionales vienen esquematizados con A ó B; las que poseen un caracter $\neq 1$ bajo la rotación principal C_n se denominan A: las que -1 . B. Las representaciones bidimensionales se designan con E y las tridimensionales con T. Cuando hay presente una inversión, se acostumbra a poner un subíndice g ó u según que sea par o impar a ella respectivamente.

Además, se suele poner al lado de cada representación irreducible, una lista de las coordenadas, las formas cuadráticas de las coordenadas y las rotaciones R_x R_y R_z según las respectivos ejes, de acuerdo a qué representación se transformen.

Como ejemplo consideremos el grupo C_{2v} cuya tabla es

C_{2v}			E	C_2	σ_v	σ'_v
x^2, y^2, z^2	Z	A_1	1	1	1	1
x_y	R_2	A_2	1	1	-1	-1
x_z	$R_{y,x}$	B_1	1	-1	1	-1
y_z	$R_{x,y}$	B_2	1	-1	-1	1

En nuestra rotación σ_v es una reflexión en el plano xz y σ'_v en el y z. Los efectos de cada una de sus operaciones sobre las tres coordenadas es

$$P_E \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad P_{C_2} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -x \\ -y \\ z \end{pmatrix} \quad P_{\sigma_v} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ -y \\ z \end{pmatrix} \quad P_{\sigma'_v} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -x \\ y \\ z \end{pmatrix}$$

vemos que las coordenadas x, y, z se transforman así según las representaciones B_1 , B_2 y A_1 respectivamente, o formalmente

$$P_R x = G^{(B_1)}(R) x \quad P_R y = G^{(B_2)}(R) y \quad P_R z = G^{(A_1)}(R) z$$

PROPIEDADES MAGNETICA

DE LA MATERIA

1 INTRODUCCION.

Es conveniente agrupar las propiedades magnéticas de los sólidos bajo los tres siguientes encabezamientos.

- (i) diamagnetismo
- (ii) paramagnetismo
- (iii) ferromagnetismo, antiferromagnetismo y ferrimagnetismo .

Veremos aquí de una manera global las propiedades de estos tres grupos sin estudiar muy a fondo los detalles cuantitativos de los mismos. Igualmente, en la mayor parte de los casos nos limitaremos a un estudio mas o menos macroscópico del fenómeno , sin llegar a detalles del proceso microscópico. La teoria cuantica, que nos daría una idea mas real , de las propiedades magnéticas de la materia, aparte de presentar la actual una notoria complejidad, reúne la característica de que aun no está elaborada completamente por lo que no parece adecuada su exposición rigurosa aquí.

Daremos primero un breve repaso de algunos conceptos macroscópicos que luego usaremos.

Cuando una sustancia se coloca en un campo magnético \vec{H} , aparece un momento magnético \vec{M} por unidad de volumen. Para materiales isótropos \vec{M} y \vec{H} son vectores paralelos y la susceptibilidad χ definida por

$$\vec{M} = \chi \vec{H} \quad (1.1)$$

es entonces una cantidad escalar. En sustancias anisotrópicas, χ es un tensor. En el caso de que \vec{M} se refiere a una molécula gramo, se introduce la susceptibilidad molar χ_m . Todos los átomos o iones producen una contribución diamagnética a la susceptibilidad total, aunque pueda quedar enmascarada por los otros tipos, es una consecuencia del momento magnético inducido en los átomos por un campo externo. Recordemos que el diamagnetismo puede compararse con la polarización electrónica en un campo eléctrico. Existe, sin embargo, una diferencia esencial. en el caso eléctrico el momento inducido se alinea según la dirección del campo aplicado, conduciendo a una susceptibilidad eléctrica positiva en el caso magnético el momento inducido produce una susceptibilidad negativa.

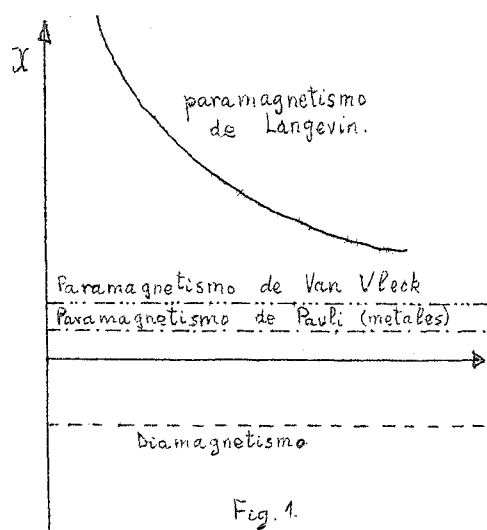
Ambos son esencialmente independientes de la temperatura.

El paramagnetismo requiere la existencia de dipolos magnéticos permanentes, y la susceptibilidad paramagnética es análoga de la susceptibilidad orientada asociada con los dipolos eléctricos permanentes. Ambos casos la susceptibilidad es positiva y dependiente de la temperatura. Las propiedades correspondientes al grupo (iii) requieren también la existencia de dipolos magnéticos permanentes y además una interacción relativamente fuerte entre ellos.

La inducción magnética \vec{B} puede ser definida como

$$\vec{B} = \mu_0 (\vec{H} + \vec{M}) = \mu_0 (1 + \chi) \vec{H} = \mu \vec{H}$$

donde μ se denomina permeabilidad. A menos que se diga lo contrario, supondremos a \vec{H} , \vec{M} y \vec{B} vectores paralelos y por ello μ será un escalar. Para materiales paramagnéticos y diamagnéticos, la permeabilidad es una constante. Para las propiedades mencionadas como (iii) la relación entre \vec{B} y \vec{H} es mucho más complicada mostrando ciclo de histéresis como veremos luego.



Las sustancias pues con susceptibilidad magnética negativa serán las que llamaremos diamagnéticos y las con positiva, paramagnéticas. La representación de las mismas en función de la temperatura puede verse en la fig. 1. Quedan allí representadas también tres tipos de paramagnetismos: los de Langevin, de Van Vleck y de Pauli para los metales.

2. DIAMAGNETISMO

2.1 ORIGEN DE LOS DIPOLOS MAGNÉTICOS PERMANENTES

Como fue postulado por la hipótesis de Ampere, los dipolos magnéticos tienen su origen en el flujo de corrientes eléctricas. De la teoría de la electricidad es bien conocido, por ejemplo, que una corriente circular estacionaria circulando en un plano produce un campo magnético que a grandes distancias puede ser descrito como resultante de un dipolo magnético.

$$\mu = I S / c \quad (2.1)$$

(Emplearemos de aquí en lo sucesivo el sistema de Gauss para nuestra formulación. La constante c que aparece en las fórmulas es la velocidad de la luz.)

es la corriente y S el area del bucle) La dirección del dipolo es perpendicular al plano del bucle.

Empleando esta relación, podemos considerar el momento del dipolo magnético asociado con un electrón que describe una órbita circular de radio r con una velocidad angular ω_o . La corriente del bucle en este caso es $-e \omega_o / 2\pi$, por lo que, de acuerdo con (2.1) el momento del dipolo magnético asociado con la órbita del electrón es

$$\mu = -e \omega_o r^2 / 2 c \quad (2.2)$$

Es interesante relacionar este momento con el angular del electrón que en este caso es $m \omega_o r^2$. Así tenemos

$$\mu = -(e/2 mc) \times \text{momento angular} \quad (2.3)$$

El signo menor indica que el momento del dipolo señala una dirección opuesta al vector que representa el momento angular.

2.2 REPASO DE LOS NUMEROS CUANTICOS

Vamos a hacer aquí un pequeño repaso de los números cuánticos que se suelen emplear en la teoría atómica. No daremos justificación de los mismo.

2.2 Número cuántico principal n : determina la energía de la órbita; puede tomar solamente valores enteros $n = 1, 2, 3, \dots$. Los correspondientes niveles electrónicos se designan con K, L, M, N, \dots

2.2,2 El momento angular de la órbita viene determinado por el número cuántico l, que queda limitado al rango de valores.

$$l = 0, 1, 2, \dots (n - 1) \quad (2.4)$$

El momento angular total asociado con un valor dado de l es

$$\hbar \left[l (l + 1) \right]^{\frac{1}{2}} \quad (2.5)$$

Los electrones asociados con los estados $l = 0, 1, 2, 3, \dots$ se denominan, respectivamente, s, p, d, f, g, \dots . Hay que señalar que los electrones en un estado S tienen siempre momento angular cero y por ello un momento magnético nulo también

2.2.3 Los componentes posibles del momento angular a lo largo de alguna dirección específica (tal como la dirección de un campo magnético externo H) quedan determinadas por el número cuántico magnético m_l - donde m queda restringido al conjunto de valores

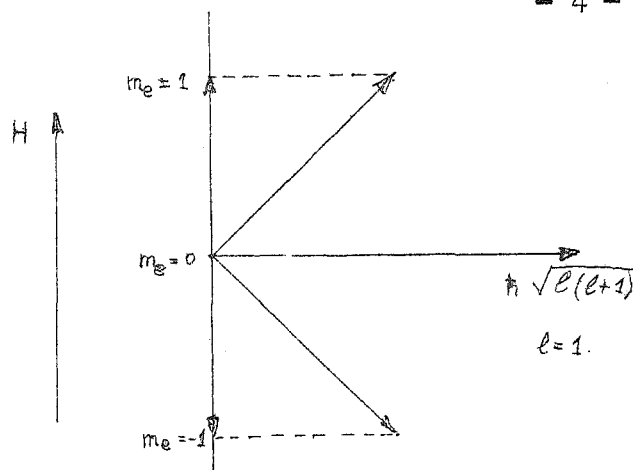


Fig. 2.1

$$m_l = l, (l-1), \dots, -l, -(l-1), \dots, -l \quad (2.6)$$

Si por ejemplo, un electrón p tiene las componentes posibles del momento angular a lo largo de la dirección de un campo magnético $\hbar, 0, -\hbar$. Consecuentemente, las componentes posibles del momento magnético a lo largo de la dirección de un campo magnético aplicado son (Fig. 2.1)

$$-e \hbar / 2mc, 0, e \hbar / 2mc$$

La cantidad $e \hbar / 2 mc = 0,927 \cdot 10^{-20}$ erg/oersted es llamada el magnetón de Bohr y será designada por μ_B .

2.2.4 Hasta aquí hemos descrito al electrón simplemente como una partícula de carga e y masa m . Sin embargo el electrón mismo tiene un momento angular conocido como spin. Las posibles componentes del momento angular del spin a lo largo de la dirección de un campo externo son $\pm \hbar/2$. Esto ha conducido a la introducción del número cuántico de spin. $s = \pm 1/2$

En la base de (2.2) debía esperarse que el spin del electrón daría lugar a una componente de medio magnetón de Bohn. Debe recalcar, sin embargo, que para el spin la relación (2.2) no es válida. De hecho la componente del momento magnético μ_{s2} del spin a lo largo de su campo externo viene dado por

$$\mu_{s2} = g (e/2 mc) (\hbar/2) \quad (2.7)$$

donde g se denomina factor de descomposición espectroscópico o razón giromagnética. Para el spin del electrón $g = 2,0023$, esto es, el spin del electrón

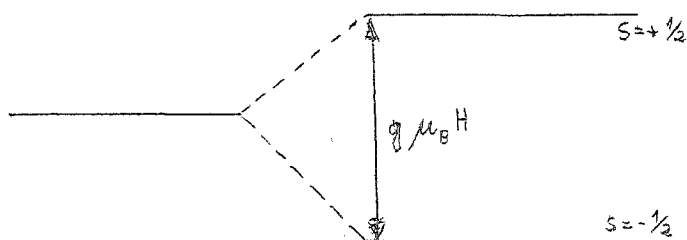


Fig. 2.2

da lugar a muy proximately un magnetón de Bohr en la dirección (u compuesta) de un campo externo H. La razón para el nombre de "factor de descomposición" es la siguiente. Consideremos un electron con spin 1/2 y sin momento angular orbital, bajo la influencia del campo magnético H. Como queda explicitado en la fig. 2.2, esto da lugar a dos niveles de energía separados por una energía

$$\Delta E = 2 \left| \mu_{s2} \right| H = g (e/2 mc) \hbar H = g \mu_B H \quad (2.8)$$

Por ello g determina el intervalo en el que el nivel original se ha desdoblado

2.2.5 El momento angular orbital y el spin pueden combinarse vectorialmente para dar el momento angular total; este último se determina por el número cuántico j. Por ello, para un electrón con un cierto l y un spin de 1/2, j solo puede tomar los valores l + 1/2. En átomos conteniendo un cierto número de electrones, los vectores l pueden combinarse para dar una resultante L, y los vectores S para dar otra resultante S. Este tipo de combinación se denomina acoplo Russell-Saunders; es el único tipo que consideraremos aquí. Las resultantes L y S se combinan entonces para formar el momento angular total J del sistema electrónico total del átomo. Para tales átomos, el factor de descomposición espectroscópico viene dado por la fórmula de Landé

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2 J (J+1)} \quad (2.9)$$

Reglas de Hund

A fin de predecir el momento dipolar asociado con el sistema electrónico de un átomo dado, las consideraciones anteriores deben combinarse con el principio de Pauli y las reglas de Hund.

Según el principio de Pauli, solamente un electron puede ocupar un estado definido por el conjunto de números cuánticos n, l, m_l y S. Se puede ver fácilmente que esto conduce de forma directa a la conclusión de que las capas electrónicas llenas no contribuyen al momento magnético de un átomo. Por ello el momento magnético en los átomos resulta de las capas incompletamente llenas. Mirando esto último, las reglas de Hund establecen que para el estado básico de tales átomos:

(i) Los spines de electrón se suman para dar el máximo S posible de acuerdo con el principio de Pauli.

(ii) Los momentos orbitales se combinan para dar el máximo - valor para L que sea consistente con (i)

(iii) Para una capa incompletamente llena

$J = L - S$ para una capa menos de la mitad ocupada

$J = L + S$ para una capa mas de la mitad ocupada

Como ejemplo, consideremos el ión C_r^{2+} , con una configuración electrónica $1 S^2, 2 S^2, 2 p^6, 3 S^2, 3 p^6, 3 d^4$. Todas las capas están llenas excepto la $3 d$, que contiene 4 electrones. Para una capa d , $l = 2$, por lo que de acuerdo con (2.6) m_l tiene $2l + 1 = 5$ valores posibles, Cada uno de estos puede acomodar 2 electrones ($S = \pm 1/2$), por lo que el máximo número de electrones en el nivel $3 d$ es 10. En el C_r^{2+} , la capa $3 d$ está por ello menos de la mitad ocupada. De acuerdo con la regla de Hund (i), tenemos $S = 2$. Los valores posibles de m_l son $+2, +1, 0, -1, -2$. Si colocamos los cuatro electrones todos con un spin $+1/2$ en los primeros cuatro de estos, obtenemos $L = 2$, que es el valor máximo consistente con la distribución de spin. De aquí, de acuerdo con (iii), tenemos en este caso $J = 0$.

Otros átomos o iones pueden tratarse de forma similar y de los valores S, L y J , el momento puede calcularse de (2.9) y (2.7),

Momentos magnéticos nucleares

Hasta aquí hemos mencionado solo el movimiento orbital y el spin de los electrones como posibles contribuidores al momento magnético del átomo. Otra contribución puede proceder del momento magnético nuclear. Este último se expresa en magnetones nucleares, en analogía con el magnetón de Bohr, definidos por

$$\mu_n = e \hbar / 2 M_p c = 5,05 \cdot 10^{-24} \text{ erg/oersted} \quad (2.10)$$

donde M_p representa la masa de un protón, Así, los momentos magnéticos nucleares son mas pequeños que los asociados con los electrones en un factor $\sim 10^3$. Los momentos magnéticos nucleares son un resultado del momento angular nuclear (spin nuclear).

Sumarizando vemos que los dipolos magnéticos se originan de :

(a) el movimiento orbital de los electrones

- (b) el spin del electrón.
- (c) el spin nuclear.

2.3 DIAMAGNETISMO Y PRECESION DE LARMOR

El principio básico del comportamiento diamagnético puede ser ilustrado facilmente con referencia a la bien conocida ley de Lenz de la Electricidad.

Consideremos una corriente cerrada con un campo magnético asociado. Cuando se intenta cambiar el flujo magnético encerrado por el bucle aplicando un campo magnético externo H , se induce una corriente de dirección tal que el campo magnético resultante de la corriente inducida contrarresta al campo H . Supongamos ahora que la resistencia eléctrica del bucle es nula; la corriente inducida persistirá entonces tanto tiempo como esté presente el campo externo. Tal situación se cumple en el bucle de corriente asociado con el movimiento de un electrón en un átomo. Consecuentemente, cualquier órbita atómica producirá una contribución negativa a la susceptibilidad magnética.

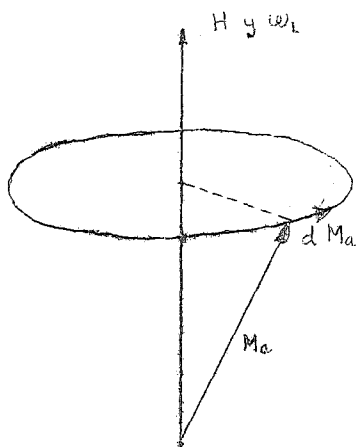


Fig. 2.3

La precesión de Larmor

Consideremos la influencia de un campo magnético sobre el movimiento de un electrón en un átomo cuantitativamente. Con referencia a la fig. 3.1, asumiremos una dirección arbitraria para el vector momento angular M_a , relativo, al campo H .

El momento dipolar magnético es, - de acuerdo con (2.3)

$$\mu = - (e/2 mc) M_a \quad (2.11)$$

El campo magnético produce un par $\mu \times H$ en el dipolo, por lo que, de acuerdo con la Mecánica Newtoniana, podemos escribir.

$$(d/dt) M_a = \mu \times H = - (e/2mc) H_a \times M \quad (2.12)$$

Esta es la ecuación de movimiento de un vector M_a en precesión en torno a H con frecuencia angular

$$\omega_L = e H/2 mc \quad (2.13)$$

donde ω_L es la frecuencia de Larmor. Esto puede verse tambien de la fig. 3.1, en la que para una precesión del tipo (3.3)

$$dM_a = \bar{\omega}_L \times M_a dt = (\bar{e}/2mc) H \times M_a dt$$

de acuerdo con (2.12), Vemos que $e/2 mc = 1,4010^6 \text{ S}^{-1} \text{ gauss}^{-1}$, por lo que aun para un campo de 10^5 gauss la frecuencia de Larmor ω_L es mucho mas pequeña - que la obtención de (2.13) se basa en la suposición de que M_a es independiente de H , esto es, se supone que la órbita no se deforma bajo la influencia - del campo magnético. Para una primera aproximación, esto es correcto.

De lo que hemos dicho hasta aquí. se infiere que bajo la - influencia de un campo externo, el plano de la órbita no es estacionario, sino que precece en torno a H . Como un resultado de la carga del electrón, la precesión produce un momento magnético inducido con una componente opuesta a la de H . De hecho, de acuerdo con (2.3), esta componente es igual a

$$(\mu_{inf})_H = - (e/2 mc) m \omega_L \langle q^2 \rangle = - (e^2/4 mc^2) H \langle q^2 \rangle \quad (2.14)$$

donde $\langle q^2 \rangle$ es el radio cuadrático medio de la proyección de la órbita sobre un plano perpendicular a H . Cuando este tratamiento se extiende a un sólido contenido N átomos por cm^3 , cada átomo conteniendo Z electrones, se obtiene para la susceptibilidad diamagnética definida como el momento inducido por cm^3 por gauss.

$$\chi_{dis} = - N Z (e^2/6 mc^2) \langle r^2 \rangle \quad (2.15)$$

Aquí hemos supuesto que la distribución de la carga de los átomos tiene simetría esférica, por lo que $\langle r^2 \rangle = \frac{3}{2} \langle q^2 \rangle$ representa la distancia cuadrática media de los electrones al nucleo.

4 (Para una distribución esférica de carga

$$\langle x^2 \rangle = \langle y^2 \rangle = \langle z^2 \rangle$$

y es v

$$\langle r^2 \rangle = \langle x^2 \rangle + \langle y^2 \rangle + \langle z^2 \rangle \quad y \quad \langle q^2 \rangle = \langle x^2 \rangle + \langle y^2 \rangle$$

La susceptibilidad diamagnética está por ello determinada esencialmente por la distribución de carga en los átomos. Notar que χ es negativa. Con $\langle r^2 \rangle = 10^{-16} \text{ cm}^2$ y con $N \approx 5 \cdot 10^{22} \text{ cm}^{-3}$, se obtiene

$$\chi \approx 10^{-7} \quad Z \approx 10^{-6}$$

3. PARAMAGNETISMO

El paramagnetismo electrónico (contribución positiva a χ) se encuentra en :

(a) Todos los átomos y moléculas que poseen un número impar de electrones pues en este caso el spin total del sistema no puede ser nulo. Ejemplos : los átomos libres de radio, el óxido nítrico gaseoso (NO), los radicales libres orgánicos tales como el trifenilmetilo $C(C_6H_5)_3$.

(b) Todos los átomos y iones libres con una capa interna in completa; elementos de transición ; los iones isoelectrónicos con elementos de transición; las tierras raras y los elementos actínidos. Ejemplos: M_n^{2+} , G_d^{3+} , U^{4+} . Las propiedades paramagnéticas se observan en muchos de estos iones cuando se incorporan a los sólidos y como iones en disolución , pero no invariablemente.

(c) Un corto número de compuestos diversos con un número par de electrones incluyendo el oxígeno molecular y los radicales dobles orgánicos

(d) Los metales.

3.1 TEORIA CLASICA DEL PARAMAGNETISMO

Consideremos un medio constituido por N átomos por unidad de volumen llevando cada uno un momento magnético dipolar μ . Supongamos que la interacción entre los dipolos es débil, por lo que el campo en el que el un dipolo dado se encuentra es igual al campo aplicado H . Supondremos aquí que el campo magnético es constante o varía lentamente con el tiempo. En la teoría clásica se supone que los dipolos rotan libremente. La imanación resulta de la orientación de los momentos magnéticos por la acción de un campo aplicado; el desorden térmico se opone a la Tendencia del campo a orientar los momentos. La energía de interacción con un campo magnético aplicado es:

$$E = - \mu \cdot \vec{H} = - \mu H \cos \theta \quad (3-1)$$

en donde θ es el ángulo entre el momento y la dirección del campo. La imanación será

$$M = N \mu \langle \cos \theta \rangle \quad (2-3)$$

siendo N el número de moléculas por unidad de volumen y $\langle \cos \theta \rangle$ el valor medio de una distribución en equilibrio térmico.

Según la ley de distribución de Boltzmann, la probabilidad relativa de encontrar una molecula en un elemento de ángulo sólido $d\Omega$ es - proporcional a $e^{-E/KT}$ y por ello

$$\langle \cos \vartheta \rangle = \frac{\int e^{-E/KT} \cos \vartheta d\Omega}{\int e^{-E/KT} d\Omega} \quad (3.3)$$

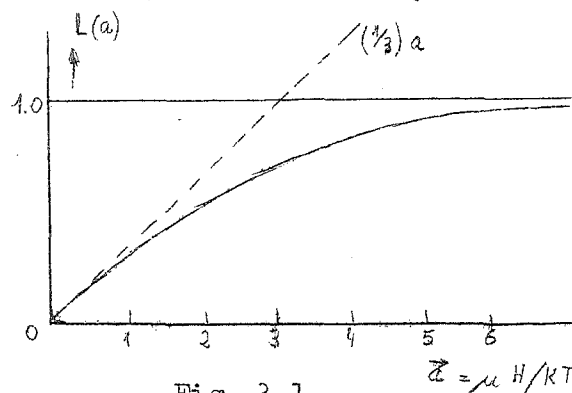
la integración debe extenderse a todos los ángulos solidos posibles, o sea.

$$\langle \cos \vartheta \rangle = \frac{\int_0^\pi 2\pi \sin \vartheta \cos \vartheta e^{\mu H \cos \vartheta / KT} d\vartheta}{\int_0^\pi 2\pi \sin \vartheta e^{\mu H \cos \vartheta / KT} d\vartheta}$$

haciendo $x = \cos \vartheta$ y $a = \mu H/KT$ queda

$$\langle \cos \vartheta \rangle = \frac{\int_{-1}^1 e^{ax} x dx}{\int_{-1}^1 e^{ax} dx} = \frac{d}{da} \frac{1}{a} \int_{-1}^1 e^{ax} dx = \text{ctha} - \frac{1}{a} \equiv L(a) \quad (3-4)$$

La función $L(a)$ se denomina la función de Langevin. En la fig. 3.1. ha sido representada $L(a)$ en función de $a = \mu H/KT$. Se ve que para altos - valores de a , esto es, para grandes valores del campo, la función se aproxima a la unidad.



Con lo anterior

se obtiene finalmente

$$M = N\mu L(a) \quad (3-5)$$

Para valores $\mu H \ll KT$ ($a \ll 1$), $L(a) = a/3$, con lo que

$$M \approx N \mu^2 H / 3KT \quad \text{ó} \quad \chi = N \mu^2 / 3KT \quad (3-6)$$

Puede verse que μ es del orden de un magnetón de Bohr $\approx 10^{-20}$ erg/gauss, por lo que para un campo de 10^4 gauss, $\mu H \approx 10^{-16}$ erg. A la temperatura ambiente $KT/3 \approx 10^{-14}$ erg, por lo que la condición $\mu H \ll KT$ se satisface excepto para muy bajas temperaturas. La relación $\chi = \text{cte}/T$ se conoce como ley de Curie.

3.2 TEORIA CUANTICA DEL PARAMAGNETISMO

De acuerdo con la teoria cuántica, el momento magnético per-

manente de un átomo dado o ion no rota libremente, sino que queda restringido a un conjunto finito de orientaciones relativas al campo aplicado. Consideremos así un medio conteniendo N átomos por unidad de volumen, el total número cuántico del momento angular de cada átomo sea J (combina el momento angular orbital total L y el spin total S del sistema electrónico por átomo). De acuerdo con 2.1 esto da lugar a las posibles componentes del momento magnético.

$$M_j \text{ g } \mu_B \quad \text{donde} \quad M_j = J, (J-1), \dots, -(J-1), -J \quad (3-7)$$

aquí M_j es el número cuántico magnético asociado con J . La energía potencial de un dipolo magnético con una componente $M_j \text{ g } \mu_B$ según H es $-M_j \text{ g } \mu_B H$, por lo que de acuerdo con la mecánica estadística, la magnetización viene dada - por

$$M = N \frac{\sum_{-j}^{+j} M_j \text{ g } \mu_B \exp(M_j \text{ g } \mu_B H/KT)}{\sum_{-j}^{+j} \exp(M_j \text{ g } \mu_B H/KT)} \quad (3-8)$$

El coeficiente de N en la parte derecha es la media estadística de la componente del momento magnético por átomo según H .

Podemos distinguir aquí dos casos:

(i) $M_j \text{ g } \mu_B H/KT \ll 1$. Bajo estas circunstancias las exponenciales de (3-8) pueden aproximarse por $(1 \pm M_j \text{ g } \mu_B H/KT)$ y realizando las sumas se encuentra fácilmente para la susceptibilidad paramagnética.

$$\chi = M/H = \frac{g^2 J(J+1) \mu_B^2}{3KT} \quad (3-9)$$

Este resultado es idéntico al resultado clásico (3-6) ya que el momento magnético total μ_j asociado con j viene dado por:

$$\mu_j^2 = g^2 J(j+1) \mu_B^2 \quad (3-10)$$

(ii) A bajas temperaturas y campos magnéticos intensos la condición dada en (i) no se satisface y (3-8) deberá calcularse sin aproximar los exponentes. Operando se llegó a:

$$M = Ng J \mu_B B_j(x) \quad (3-11)$$

donde $x = g J \mu_B H/KT$ y $B_j(x)$ es la función de Brillouin definida por

$$B_J(X) = \frac{2^{J+1}}{2^J} \operatorname{cth} \left[\frac{(2J+1) X}{2^J} \right] - \frac{1}{2^J} \operatorname{cth} \left(\frac{X}{2^J} \right) \quad (3-12)$$

Esicamente hablando, este resultado implica saturación de la magnetización a bajas temperaturas, esto es, todos los dipolos estarán dirigidos finalmente según H. En este respecto (3-11) es la análoga de la expresión de Langevin (3-5), la diferencia estriba unicamente en la notación libre de los dipolos. De hecho, si $J \rightarrow \infty$ (infinito número de orientaciones posibles), la expresión de Brillouin pasa a ser idéntica a la de Langmuir.

El orden de magnitud de la susceptibilidad paramagnética de un sólido por cm^3 puede estimarse de (3-9). Con $N \approx 10^{22}$ y un momento dipolar de un magnetón de Bohn, se obtiene $\chi \approx 1/300 \text{ T}$. A la temperatura ambiente $\chi \approx 10^{-5}$; o 1°K $\chi \approx 10^{-3} - 10^{-2}$. Estos valores son de importancia en conexión con la siguiente cuestión: en la teoría de la polarización dieléctrica de un sólido, era necesario la introducción de un campo eléctrico interno, esto es, el campo actuando sobre un átomo dado estaba representado por la suma del campo aplicado y el campo debido a la polarización de los vecinos. Mas aquí, para la anterior derivación de la susceptibilidad magnética, el campo magnético que actuaba sobre el dipolo magnético se suponía era igual solo al campo aplicado H. La justificación para esto es la siguiente: el orden de magnitud del campo interno viene dado por $H + \gamma M = H(1 + \gamma \chi)$ donde $\gamma \approx 4$. De aquí, el error funcional hecho despreciando la corrección del campo interno es del orden de χ . Como hemos visto antes, este es pequeño para materiales paramagnéticos.

Debe mencionarse finalmente que existe también una contribución paramagnética independiente de la temperatura a la susceptibilidad a bajas temperaturas.

Es el denominado paramagnetismo de Van Vleck.

4. HAMILTONIANO DE UN ELECTRON EN UN CAMPO MAGNETICO

Es interesante considerar el problema del para y el diamagnetismo desde un punto de vista diferente, partiendo de la ecuación de Lorentz para la fuerza que actúa sobre un electrón moviéndose en un campo combinado eléctrico y magnético.

$$F = -eE - \frac{e}{c} \mathbf{V} \times \mathbf{H} \quad (4-1)$$

No consideraremos aquí al spin del electrón. Habiendo definido el potencial vector magnético \vec{A} por la relación $\vec{H} = \nabla \times \vec{A}$, el hamiltoniano de un electrón en un campo magnético, es una función de la forma.

$$\mathcal{H} = \frac{(\vec{p} - e \vec{A}/c)^2}{2m} + U(x, y, z) \quad (4.2)$$

donde $U(x, y, z)$ es la energía de la partícula en el campo eléctrico externo. Elevando al cuadrado quedará

$$\mathcal{H} = \frac{1}{2m} \left[\vec{p}^2 - \frac{e}{c} \vec{p} \cdot \vec{A} - \cancel{\frac{e}{c} \vec{p} \cdot \vec{A}} - \frac{e}{c} \vec{A} \cdot \vec{p} + \frac{e^2}{c^2} A^2 \right] + U(x, y, z) \quad (4.3)$$

En mecánica Cuántica, en la representación de Schrödinger, el momento \vec{P} es el operador $-i\hbar \nabla$; con ello el hamiltoniano resultante será el de la partícula libre mas unos sumandos debidos a la presencia del campo magnético. Estos sumandos son:

$$\mathcal{H}' = \frac{ie\hbar}{2mc} (\nabla \cdot \vec{A} + \vec{A} \cdot \nabla) + \frac{e^2}{2mc^2} A^2 \quad (4.4)$$

que pueden tratarse como una pequeña perturbación. Si suponemos a la componente z en la dirección del campo magnético, esto es

$$H_x = H_y = 0 \quad \text{y} \quad H_z = H \quad (4.5)$$

quedará

$$A_x = -\frac{1}{2} y H \quad A_y = \frac{1}{2} x H \quad A_z = 0 \quad (4.6)$$

con lo que (4.4) se convierte en

$$\mathcal{H}' = \frac{ie\hbar H}{2mc} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) + \frac{e^2 H^2}{8mc^2} (x^2 + y^2) \quad (4.7)$$

Ahora bien, de la definición de momento angular

$$\vec{M}_a = \vec{r} \times \vec{p}$$

se infiere que el primer sumando resulta proporcional a su componente Z

$$x p_y - y p_x = (M_a)_z$$

Y como por otra parte, si el movimiento del electrón estuviera asociado con un momento dipolar magnético permanente μ , contribuiría con un término $-\mu \cdot \vec{H} = -\mu_z H$ en el Hamiltoniano, el primer término puede ser identificado también con $\mu_z H$, por lo que

$$\mu_z = -\frac{e\hbar}{2mc} (x P_y - y P_x)$$

y recordando la relación (2.3), se llega a que la relación

$$\mu = -\frac{e\hbar}{2mc} \times \text{momento angular}$$

entre el momento dipolar permanente y el momento orbital angular se obtiene de forma directa del Hamiltoniano.

En segundo lugar, si nos fijamos en el término de H^2 en la expresión (4.7) si hubiéramos expresado el Hamiltoniano para los electrones asociados con un volumen unidad de una sustancia conteniendo N átomos y cada uno con Z electrones, el término en H^2 sería .

$$N (e^2/8 mc^2) H^2 \sum_{i=1}^N (x_i^2 + y_i^2) = NZ (e^2/8 mc^2) H^2 \langle r^2 \rangle \quad (4.8)$$

donde $\langle r^2 \rangle$ representa la media de los cuadrados de los radios de las proyecciones de las orbitas en un plano perpendicular a H . Ahora, si el campo magnético induce un momento de dipolo en el material, el correspondiente término de la energía debería ser cuadrático en H . Por ello (4.8) debe ser considerado el término energético asociado con el diamagnetismo del sólido. Comparando (4.8) y el valor de la energía por unidad de volumen $\frac{1}{2} \chi H^2$ se obtiene

$$\chi_{\text{dia}} = -NZ (e^2/4 mc^2) \langle r^2 \rangle = -NZ (e^2/6 mc^2) \langle r^2 \rangle$$

donde $\langle r^2 \rangle$ representa la distancia cuadrática media de los electrones a los núcleos. Vemos que este resultado es idéntico al (2.15)

5. FERROMAGNETISMO.-

5.1 Consideraciones preliminares.- Los materiales ferromagnéticos son aquellos que poseen un momento magnético espontáneo, es decir, un momento magnético incluso en ausencia de un campo magnético aplicado. Definimos así una imanación de saturación M_s como el momento magnético espontáneo por unidad de volumen. Por encima de una temperatura crítica θ_f , conocida como temperatura ferromagnética de Curie, la magnetización espontánea desaparece y el material pasa a ser paramagnético. Un poco por encima de la temperatura de Curie, la susceptibilidad sigue la ley de Curie-Weiss

$$\chi = C / (T - \theta) \quad (5.1)$$

donde C es la constante de Curie; la temperatura θ se denomina temperatura paramagnética de Curie y usualmente es algunos grados mayor a θ_f .

La teoría del ferromagnetismo se centra en torno a las dos siguientes hipótesis establecidas en 1.907 por Weiss.

(i) Una especie ferromagnética de dimensiones macroscópicas contiene, en general, un cierto número de pequeñas regiones (dominios) que se encuentran espontáneamente magnetizados; la magnitud de la magnetización espontánea de la especie viene determinada por la suma vectorial de los momentos magnéticos de los dominios individuales.

(ii) Dentro de cada dominio la magnetización espontánea se debe a la existencia de un "campo molecular" que tiende a producir una alineación paralela de los dipolos atómicos.

Solamente unos pocos materiales muestran ferromagnetismo. Estos incluyen los elementos Fe , Ni , Co , Gd y Dy , sus aleaciones entre ellos y con algunos otros elementos, y unas pocas sustancias no conteniendo ninguno de los elementos antes citados. La gran mayoría de los materiales ferromagnéticos son metales o aleaciones; unos pocos, sin embargo son compuestos iónicos, tales como $CrBr_3$, EuO , EuS , $EuSe$, EuI_2 , y $EuSiO_4$. Los materiales ferrimagnéticos también presentan magnetización espontánea; son los que veremos más adelante.

Los elementos Tb, Ho, Er y Tm son también ferromagnéticos a bajas temperaturas.

Los dos fenómenos establecidos en las hipótesis de Weiss llevan consigo la división del estudio del ferromagnetismo en dos grandes áreas. La primera se refiere esencialmente al origen y propiedades de la magnetización espontánea dentro de un dominio aislado. La segunda se refiere al comportamiento de la magnetización de una especie con un campo magnético aplicado. La especie consistirá, en general, de varios dominios; sin embargo, bajo ciertas condiciones, podrá ser uno solo.

Estudiaremos únicamente la teoría clásica de la magnetización espontánea desarrollada por Weiss. En esta teoría fenomenológica la interacción fuerte entre los dipolos atómicos, denominada campo molecular o de Weiss, se supone es proporcional a la magnetización. El origen de este campo reside en las fuerzas de intercambio de la mecánica cuántica y queda fuera de nuestro fin actual. Aunque la teoría de Weiss es satisfactoria en los puntos principales existen unas ciertas diferencias en los detalles, diferencias que se corrigen con otros modelos más reales.

5.2. Campo molecular de Weiss.- La magnetización espontánea implica una cooperación entre los dipolos atómicos dentro de un dominio simple, esto es, debe existir una especie de interacción entre los átomos que produce la tendencia para la alineación paralela de los dipolos atómicos magnéticos. A fin de obtener una descripción fenomenológica de la magnetización espontánea, Weiss supuso que el campo molecular H_m actuante sobre un dipolo dado debe ser escrito en la forma

$$H_m = H + \gamma M \quad (5.2)$$

donde H es el campo externo aplicado, M es la magnetización y γ es la constante de Weiss. Claramente se ve que el término γM da el efecto cooperativo. Sin dar una interpretación física de la constante γ , mostraremos aquí que un campo de la forma (5.2) conduce a la magnetización espontánea, a la existencia del punto ferromagnético de Curie y a la ley de Curie-Weiss (5.1). Usaremos la teoría cuántica de la magnetiza

ción mejor que la teoría clásica de Langevin usada por Weiss en su artículo original.

Consideremos un sólido conteniendo N átomos por unidad de volumen cada uno de los cuales con número cuántico de momento angular total J (que incluye la contribución orbital total L y la contribución de spin total S). De acuerdo con los resultados vistos en 3, se podrá escribir para la magnetización.

$$M = Ng \mu_B J B_J(x) \quad (5.3)$$

donde para sólidos paramagnéticos $x = g \mu_B H J / KT$. Para materiales ferromagnéticos reemplazaríamos H por H_m , de acuerdo con la suposición (5.2), ya que es H_m el campo actualmente visto por un dipolo atómico dado. Así, en el caso actual

$$x = g \mu_B (H + \gamma M) J / KT \quad (5.4)$$

ya que lo que nos interesa es la magnetización espontánea, será $H = 0$ y así

$$M = x K T / \gamma g \mu_B J \quad (5.5)$$

Como M debe satisfacer (5.3) y (5.5), su valor a una temperatura dada deberá obtenerse del punto de intersección de los dos correspondientes M en función de x , como queda indicado en la fig. 5.1. Notar que (5.5) es una recta de pendiente proporcional a T . De la fig. 5.1 se ve que para $T < \theta_f$ se obtiene un valor no nulo para M , aunque sea nulo, $H=0$, el campo externo aplicado. De aquí que para $T < \theta_f$ tengamos magnetización espontánea. Para $T = \theta_f$, la pendiente de la recta representada por (5.5) es igual a la de la tangente de la curva (5.3) en el origen. Por ello, para $T > \theta_f$ la magnetización espontánea desaparece.

Es evidente, por otra parte, que deberá existir una relación entre la temperatura de Curie θ_f y la constante del

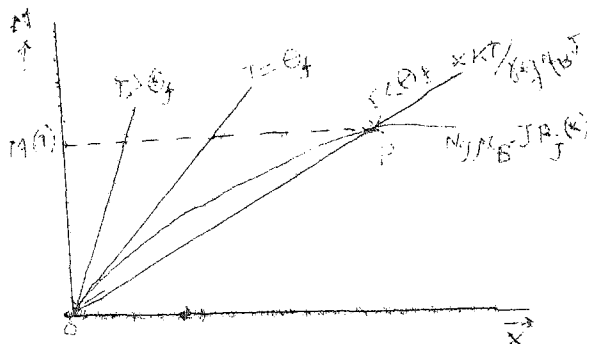


FIG.5.1

campo molecular χ ; de hecho, se espera que θ_f incremente con χ ya que la tendencia para un alineamiento paralelo crece para mayores χ . A fin de establecer esta relación, hacemos uso del hecho de que para $x \ll 1$ (cerca del origen en la fig. 5.1), la función de Brillouin viene dada aproximadamente por

$$B_j(x) \approx (J + 1) x / 3 J \quad x \ll 1 \quad (5.6)$$

de aquí que la tangente de la curva (5.3) en el origen tenga una pendiente igual a $N_g \mu_B (J + 1)/3$. Haciendo esta igual a la de (5.5) para $T = \theta_f$, se obtiene

$$3 K \theta_f / \chi = N_g^2 \mu_B^2 J(J + 1) = N \mu^2 \quad (5.7)$$

donde μ es el momento magnético total por átomo. De aquí que θ_f es proporcional a la constante del campo molecular.

Consideremos ahora la susceptibilidad en la región por encima de la temperatura ferromagnética de Curie. En esta región la magnetización ocurre solamente cuando se aplica un campo externo H ya que no hay magnetización espontánea. Por ello, para campos suficientemente pequeños con los que nos encontramos lejos de la saturación, podemos aproximar la aproximación (5.6) para $B_j(x)$ y (5.3) pasa a ser

$$M = N_g \mu_B (J + 1) x / 3 \quad (5.8)$$

donde x viene dada por (5.4). Despejando M/H después de sustituir x en (5.8) se obtiene fácilmente la ley de Curie-Weiss.

$$\chi = M/H = C/(T - \theta) \quad (5.9)$$

donde $C = N \mu^2 / 3K$ y $\theta = \chi N \mu^2 / 3K = \chi C$. Notar que el valor obtenido aquí para θ es idéntico con el obtenido para θ_f de la expresión (5.7). En otras palabras, la teoría de Weiss no distingue entre las temperaturas ferro y paramagnéticas de Curie.

5.3. COMPARACION DE LA TEORIA DE WEISS CON LA PRACTICA

5.3.1. Dependencia de la temperatura de la magnetización espontánea.

La componente máxima de un dipolo atómico asociada con un número cuántico J en cualquier dirección dada es $g J \mu_B$. De aquí el valor má-

ximo de la magnetización espontánea viene dada por $N_g / \mu_B J$, donde N es el número de átomos por unidad de volumen. Esto se infiere también de (5.3) ya que para $x \rightarrow \infty$ la función de Brillouin $B_j(x) \rightarrow 1$. De acuerdo con la fig. 5.1 la magnetización espontánea máxima ocurre para $T = 0$ y escribiremos por ello $N_g / \mu_B J = M(0)$. A fin de describir la dependencia con la temperatura de la magnetización espontánea de una forma conveniente, reescribiremos (5.3) en la forma

$$M(T)/M(0) = B_j(x) \quad (5.10)$$

donde $M(T)$ es la magnetización a la temperatura T . Similarmente, podemos escribir (5.5) en la forma

$$M(T)/M(0) = xKT / \gamma N_g^2 / \mu_B^2 J^2 =$$

$$= x T (J + 1) / 3 J \theta_f \quad (5.11)$$

donde la última igualdad se obtiene sustituyendo en función de θ_f mediante (5.7). La cantidad $M(T)/M(0)$ debe satisfacer (5.10) y (5.11); de aquí que pueda obtenerse por el método

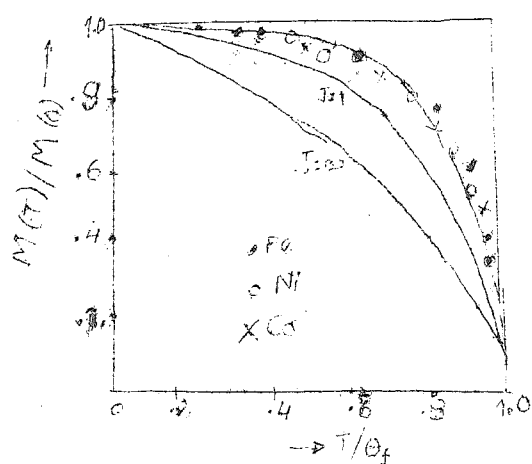


FIG.5-2

de intersección dado antes. Es importante notar que para un valor dado de J , este procedimiento conduce a una curva universal cuando $M(T)/M(0)$ se dibuja en función de T/θ_f . En la figura 5.2 se han representado tales curvas para $J = \frac{1}{2}, \infty$. El último caso corresponde a los dipolos rotando libremente según la teoría clásica. En la misma figura se han representado los puntos experimentales para el Fe, el Ni y el Co. Se observa que la curva para $J = \frac{1}{2}$ se ajusta mejor a la realidad lo que indica que la magnetización está asociada esencialmente con el spin del electrón mas que con el momento orbital de los electrones.

5.3.2 Región paramagnética.- Estudios experimentales del comportamiento del Fe, Ni y Co por encima de sus puntos de Curie han sido realizados por Sucksmith, Pearce y Fallot. De acuerdo con la ley de Curie-Weiss (5.9), una representación de $1/\chi$ en función de T daría una línea recta que cortaría al eje de T en un punto igual a 0. Los experimentos

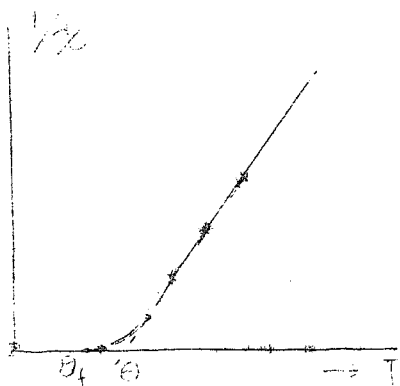


FIG. 5.3

muestran que esta ley se satisface con mucha fidelidad excepto en una región cercana al punto de Curie. De hecho; en los tres metales existe una cierta concavidad cerca del punto de Curie que dá lugar a la distinción entre las temperatu-

ras paramagnética Θ y Ferromagnética Θ_f de Curie. Este comportamiento viene indicado esquemáticamente en la fig. 5.3. Para ilustrar esto, damos a continuación Θ_f y Θ en grados Kelvin para estos metales.

	Fe	Co	Ni
Θ_f	1.043	1.393	631
Θ	1.093	1.428	650

5.4. Consideraciones cualitativas en torno a los dominios. A temperaturas inferiores al punto de Curie, los momentos magnéticos electrónicos de una especie ferromagnética están esencialmente alineados si se observan a escala microscópica. Sin embargo, observando la especie globalmente, el momento total puede ser inferior al correspondiente a la saturación y puede ser necesario la aplicación de un campo magnético externo para saturar la muestra. El comportamiento de los monocristales es en este aspecto semejante al de las especies policristalinas.

Weiss explicó este fenómeno suponiendo que las sustancias reales están compuestas de cierto número de pequeñas regiones llamadas dominios, en el interior de las cuales la imanación local está saturada; las direcciones de los diferentes dominios, sin embargo, no son necesariamente paralelas. En la fig. 5.4 se muestra

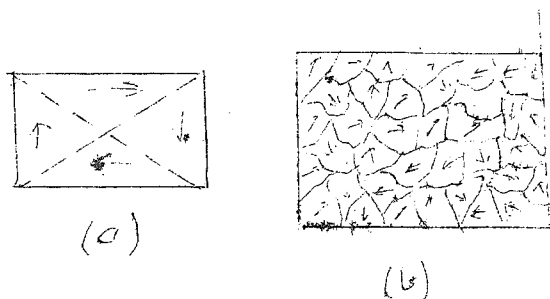


FIG. 5.4

una distribución esquemática de los dominios de un monocristal (a) y de un policristal (b), ambos con momento magnético resultante nulo.

De acuerdo con la teoría de los dominios puede imaginarse que el

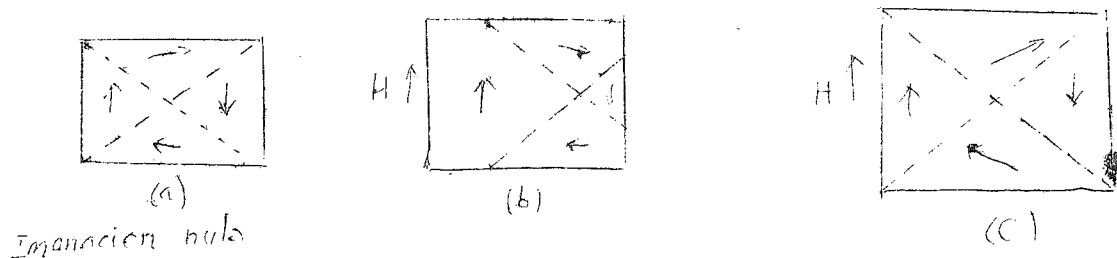


FIG. 5.5

incremento del momento magnético resultante de la especie bajo la acción de un campo magnético aplicado tiene lugar según dos procesos independientes, según sugirió R. Becker: por un aumento del volumen de los dominios que se orientaron favorablemente con respecto al campo a expensas de los dominios mal orientados (fig.5.5.b) o por una rotación de las direcciones de imanación en el sentido del campo (fig.5.5c).

Un examen detenido muestra que en campos débiles las variaciones de imanación son debidas a desplazamientos de los límites de los dominios de modo que cambia el tamaño de estos.

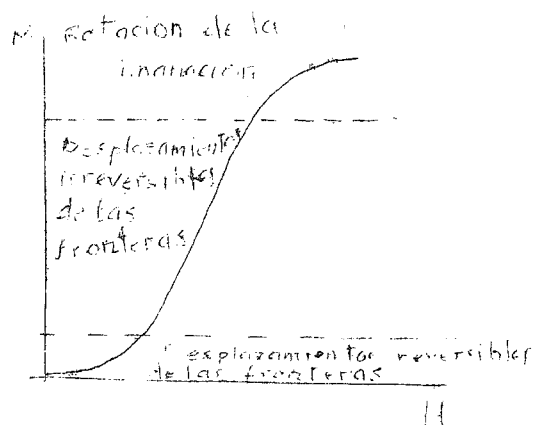


FIG. 5:6

En campos intensos la imanación cambia por medio de una rotación de la dirección en que se ejerce. En la fig. 5.6 se muestra una curva típica de imanación indicándose las regiones en que es dominante cada proceso.

La evidencia experimental más directa de la existencia de los dominios viene dada con la técnica de los diagramas de polvos magnéticos de Bitter. Este método consiste en situar una gota de suspensión coloidal de un material ferromagnético finalmente dividido, tal como la magnetita, sobre la superficie cuidadosamente preparada del cristal ferromagnético que ha

de estudiarse. La observación microscópica muestra que las partículas coloidales en suspensión se concentran intensamente a lo largo de ciertas líneas bien definidas que representan los límites entre los dominios imantados en direcciones diferentes. Ello es debido a la existencia de campos magnéticos locales muy intensos que atraen las partículas magnéticas.

El origen de los dominios puede entenderse considerando las estructuras indicadas en la fig. 5.7, cada una de las cuales representa la sección transversal de un monocristal ferromagnético. En (a) tenemos una con

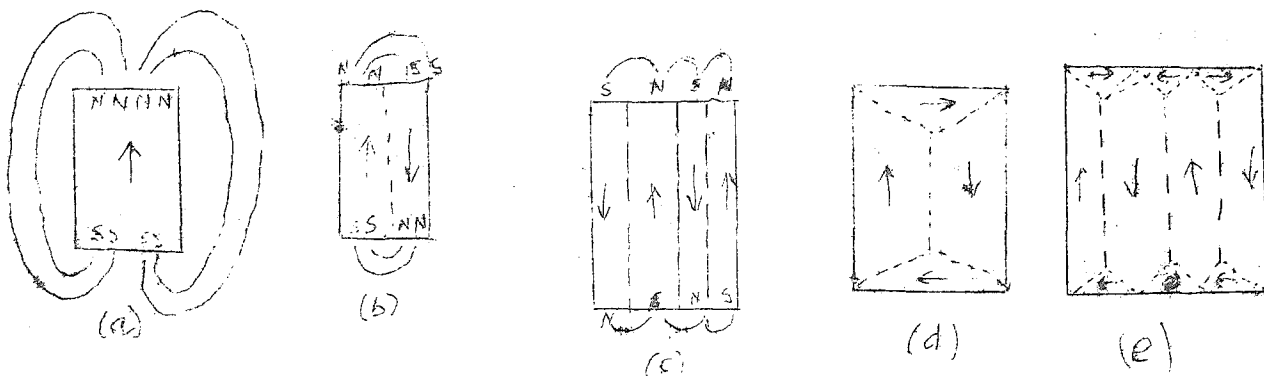


FIG. 5.7

figuración saturada formada por un solo dominio; a causa de los "polos" magnéticos formados en las superficies del cristal esta configuración poseerá un valor elevado de la energía magnética. $(1/8\pi) \int H^2 dv$.

En (b) la energía magnética se ha reducido en un factor aproximadamente igual a $1/2$ como resultado de dividir el cristal en los dominios imantados en sentidos opuestos. El proceso de subdivisión puede continuarse como en (c): con N dominios la energía magnética se reduce (a causa de la extensión espacial reducida del campo) aproximadamente a $1/N$ de la energía magnética de la configuración saturada (a).

El proceso de subdivisión prosigue hasta que la energía requerida para establecer una nueva capa de transición o interfase, separando los dominios imantados en sentidos contrarios, es superior a la reducción de la energía del campo magnético que resultaría de una subdivisión más fina. Una capa de transición va asociada a cierta cantidad de energía: en

lados opuestos del límite de imanación está dirigida en direcciones anti paralelas; como las fuerzas de intercambio favorecen las orientaciones paralelas de la imanación y se oponen a las antiparalelas, habrá que suministrar cierta energía para establecer una capa de separación.

Es posible imaginar distribuciones de dominios tales como (d) y (e) para las cuales la energía magnética es cero. En estos casos los límites de los dominios en forma de prisma triangular (llamados dominios de cierre) próximos a las caras terminales del cristal, forman ángulos iguales (45°) con la imanación de los dominios rectangulares y con la imanación de los dominios de cierre: por tanto la componente de la imanación normal a la capa de transición es continua a través del límite y no se forman polos en parte alguna del cristal. Al no existir polos no existirá campo magnético asociado a la imanación y puede decirse que el flujo se completa en circuito cerrado dentro del cristal.

5.5 La pared de Bloch.— El término " pared de Bloch" designa la capa de transición que separa los dominios adyacentes imanados en direcciones diferentes.

La idea esencial de la pared de Bloch es que la variación total de la dirección de Spin entre los dominios imanados en direcciones diferentes no se produce de un modo discontinuo en un simple plano atómico. En su lugar, el cambio de dirección tendría lugar de un modo gradual en varios planos atómicos (fig.5.8). La naturaleza gradual del cambio se debe al hecho de que para una variación total de la dirección de Spin, la energía de intercambio es menor cuando el cambio se distribuye en muchos spins que cuando se realiza bruscamente.

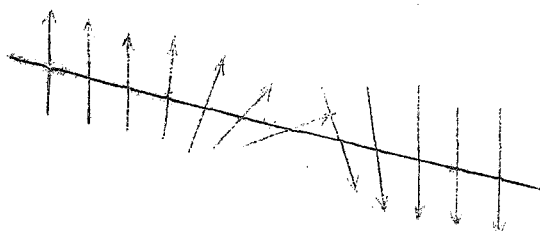


FIG. 5.8